



# Recherche systématique de quasars dans les grands sondages du ciel et application à l'étude de corrélation dans le milieu intergalactique

Franck Coppelani

## ► To cite this version:

Franck Coppelani. Recherche systématique de quasars dans les grands sondages du ciel et application à l'étude de corrélation dans le milieu intergalactique. Astrophysique [astro-ph]. Université Pierre et Marie Curie - Paris VI, 2006. Français. NNT : . tel-00127046

**HAL Id: tel-00127046**

**<https://theses.hal.science/tel-00127046>**

Submitted on 28 Jan 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Recherche systématique de quasars dans les grands sondages du ciel et application à l'étude de corrélation dans le milieu intergalactique

## Thèse

présentée et soutenue le 23 novembre 2006  
pour l'obtention du  
**Doctorat de l'Université Pierre et Marie-Curie - Paris VI**

Spécialité :  
Astronomie et Astrophysique  
par  
Franck COPPOLANI

sous la responsabilité de  
Patrick PETITJEAN et Cédric LEDOUX

### Composition du jury :

Patrick	Boissé	
Catherine	Boisson	
Rodrigo	Ibata	<i>Rapporteur</i>
Roser	Pello	<i>Rapporteur</i>
Patrick	Petitjean	<i>Directeur de thèse</i>
Laurent	Vigroux	





## Remerciements

*"Life is like topography, Hobbes. There are summits of happiness and success...flat stretches of boring routine...and valleys of frustration and failure."*

– Calvin

Je remercie tout le monde...

Voici la seule page qui ne nécessitera pas un passage par l'organe de censure qu'est mon superviseur de thèse... J'exagère bien sûr, mais ce n'est que pour mieux entamer les débuts de remerciements que je dois à Patrick Petitjean. Tout d'abord pour m'avoir accepté en tant qu'étudiant pour le stage de fin de DEA, puis pour m'avoir mis sur ce projet de thèse. Tout au long de celle-ci il m'a accompagné, guidé et conseillé. Et ce jusqu'à la fin (...) où il aura énormément apporté au texte que vous lirez. Il aura consacré une grande partie de son temps à s'assurer que cette thèse se passe bien, et je l'en remercie.

Je me dois à la fois de remercier Cédric Ledoux, mon superviseur, même s'il n'aime pas être appelé comme cela, et lui présenter mes excuses. Il m'a lui aussi accompagné et conseillé pendant les deux années que j'ai passé au Chili, mais en plus il m'a accueilli, ce dont je n'ai pas su le remercier correctement. Merci à Céline pour ses multiples relectures afin d'éliminer toutes fautes de ce mémoire (sauf de ce passage qui n'aura pas été relu).

Il est tout aussi naturel pour moi d'adresser mes remerciements à Srianand Raghunathan, ainsi qu'à son étudiant Hum Chand pour leur accueil lors de mon séjour en Inde. Je remercie aussi Jasjeet Singh Bagla pour son accueil lors de l'école à laquelle j'ai assisté à Allahabad. Merci à Danielle Alloin et à Félix Mirabel, et à travers eu l'ESO Chili, pour m'avoir accueilli et accompagné durant deux années. J'adresse également mes remerciement à Patrick Boissé pour ces conseils au long de ma thèse et pour avoir accepté de relire ce mémoire.

Enfin je remercie Roser Pelo pour m'avoir aidé au début de cette thèse et de m'avoir fourni le code remodelé d'HyperZ ainsi que pour avoir accepté d'être rapporteur de ma thèse. Merci à Rodrigo Ibata pour avoir accepté d'être rapporteur de cette thèse. Merci à Catherine Boisson pour ses commentaires sur le texte de ce mémoire qui l'ont enrichi.

Il faut dire qu'une thèse ne se résume pas uniquement aux heures passées au laboratoire, en face de son écran d'ordinateur. C'est aussi beaucoup de relations humaines, et avec les différents endroits où j'ai pu me rendre, cela fait de quoi écrire un roman. Et dans ce cas, l'importance de l'ordre dans lequel on remercie n'a que peu de signification selon moi, mais comme il faut bien commencer quelque part je classerai ces remerciements par ordre de leur apparition à l'écran de ma vie.

Merci à ma mère pour ne pas avoir trop mal pris le fait de me voir partir à l'autre bout du monde, pour son soutien et sa patience et être venu visiter le Chili avec moi.

Merci à Adlan pour m'avoir motivé à partir et ne pas m'avoir oublié malgré la distance, pour être venu jusqu'au Chili pour me rendre visite, m'avoir apporté les bon gâteaux de sa maman, et pour m'avoir longtemps fourni en musique.

Merci à Manu pour avoir fait le voyage jusqu'au Chili en sachant pertinemment

que ce ne serait pas le club Med...

Merci à Geuth pour m'avoir invité à son mariage, signe que malgré la distance je n'étais pas si loin.

Merci à mon nain préféré pour son soutien constant, pour ses bons conseils et pour avoir fait le voyage jusqu'en Amérique du Sud, ce fut une foutue aventure ! Il nous reste encore la Bolivie.

Merci à binôme pour être resté mon binôme, pour n'avoir raté aucun de mes retours, et pour les moments que l'on a réussi à partager, encore une fois malgré la distance.

Merci à Jef pour avoir été absent à chacun de mes retours en France mais pour avoir réussi à être au Pérou en même temps que moi, pour ses discussions et les soirées partagées. Et merci pour son Forum !

Merci à Gwen et Alex pour avoir réussi à organiser leur mariage en même temps que celui de Geuth, merci pour avoir été présent lors de mes passages en France.

Merci aux Nico pour les moments partagés et les moments ratés, pour les conseils de l'un et les coups de gueule de l'autre.

Merci à gugette pour les longs moments de discussion pour ses cartes et ses encouragements, merci pour avoir amené son binôme jusqu'au Chili. Merci à Pat pour avoir amené gugette jusqu'au Chili, merci pour leur visite et pour leur amitié.

Merci à Bastien pour ces conseils avisés et pour avoir toujours répondu présent lorsque j'ai eu besoin d'un coup de main.

Un très très grand merci pour Emmanuel, sa femme Elisabeth et leur deux filles Maëlys et Ondine, pour leur accueil et leur soutien lors de mon séjour en Inde. Merci pour leur soutien, et leur hospitalité.

Merci à Yuri et à Zafart pour avoir partagé tant de repas en Inde en toute circonstances !

Merci à Félix pour ses multiples coups de main.

Merci à Sylvain pour son coup ponce au Chili.

Merci au premier sourire de ma belle.

Merci à Alessandro pour m'avoir si bien accueilli et pour sa presque constante bonne humeur.

Merci à Carla pour avoir partagé MON bureau pendant presque deux années, pour tout ces instants partagés, heureux ou moins heureux, ce qui reste vaut le détour.

Merci à Stéphane pour ses longues discussions, pour son aide, ses conseils, ses coups de gueule, son attention, son humeur, bref merci d'être lui et d'avoir été là.

Merci à Linda et Klaus pour leur accueil et les bons moments passés ensemble.

Merci à Claudio pour sa bonne humeur.

Merci à Becky pour son sourire, son accueil et son sens de l'hospitalité.

Merci à Pascale pour son humeur changeante du jour au lendemain, pour ses fous rires, les moments de délire, de colère, de coup de gueule.

Merci à Hannes pour m'avoir surnommé "mighty mighty franck", merci pour les moments de complicités, ils ne tomberont pas dans le néant.

Bonne chance à Pasquier, Stefan et Kristin pour leur parcours.

Un grand merci à tout les gens de Vitacurra, de La Silla et de Paranal avec qui

j'ai partagé quelques instants, parfois pas grand chose, mais que je n'oublierai pas.

Merci à Elisabet pour m'avoir laissé sa place dans les caves de l'IAP.

Merci à Rodney pour avoir apporté un peu de soleil brésilien dans les caves de l'IAP.

Merci à Kumiko pour avoir toléré ma présence dans le bureau pendant la fin de la thèse.

Merci à Calvin et Hobbs et à leur créateur Bill Watterson pour moments de détente offert entre deux chapitres.

Merci à Brassens, Baudelaire et bien d'autre pour leur poésie.

Merci aux pisco sour, aux capirina, aux jus de fraises, framboises, pêches... à la salsa et autres compagnons de soirée Chilienne.

Merci aussi à mon laptop d'avoir tenu le temps nécessaire à cette thèse, même s'il s'en est fallu parfois de peu... (je croise encore les doigts au moment où j'écris cette ligne).

.... et merci encore à Céline, simplement, elle saura pourquoi, je compte bien lui rappeler tout les jours.



# Table des matières

<b>1</b>	<b>Introduction</b>	<b>19</b>
1.1	Le milieu intergalactique, son histoire, son étude à travers les spectres des quasars . . . . .	20
1.2	Comment rechercher les quasars ? . . . . .	24
1.2.1	Cette belle histoire qu'est la recherche de quasars . . . . .	25
	Description des quasars. . . . .	25
	Les catalogues existants de quasars . . . . .	27
1.2.2	Méthodes de recherche suivant les différents domaines d'énergie . . . . .	32
	Recherche dans le domaine optique . . . . .	33
	Recherches venues d'ailleurs . . . . .	35
<b>2</b>	<b>Les sondages utilisés : CFHTLS et SDSS</b>	<b>37</b>
2.1	SDSS . . . . .	38
2.1.1	La photométrie du SDSS . . . . .	39
2.1.2	Échantillons du SDSS . . . . .	42
	Base Spectroscopique . . . . .	44
	Base Photométrique . . . . .	45
	Catalogue de quasars . . . . .	47
2.2	CFHTLS . . . . .	47
2.2.1	Généralités - Traitement et accès aux données . . . . .	47
2.2.2	Photométrie - Comparaison avec le SDSS . . . . .	49
2.3	Autres champs observés au CFHT . . . . .	56
<b>3</b>	<b>Présentation des méthodes de sélection de quasars</b>	<b>59</b>
3.1	Les propriétés des quasars . . . . .	61
3.1.1	Spectre en émission des quasars . . . . .	61
	Domaine radio . . . . .	62
	Infra-rouge . . . . .	63
	Domaine X . . . . .	63
	Optique et ultra-violet . . . . .	64
3.1.2	Les absorptions dans les spectres des quasars . . . . .	66
	Systèmes métalliques . . . . .	66
	Systèmes de la forêt Lyman- $\alpha$ . . . . .	67
	Large raies d'absorption (BAL) . . . . .	67
3.2	Introduction à la sélection de quasars . . . . .	69

3.3	Diagrammes couleur-couleur . . . . .	71
3.3.1	Le locus des étoiles dans les diagrammes couleur-couleur .	72
3.3.2	Position des galaxies dans les diagrammes couleur-couleur	74
3.3.3	Position des quasars dans les diagrammes couleur-couleur	75
3.3.4	Algorithmes de sélection des quasars dans le SDSS . . . . .	79
	Premier algorithme de sélection du SDSS . . . . .	79
	Seconde sélection du SDSS . . . . .	86
	Coupe dans les diagrammes couleur-couleur . . . . .	88
3.3.5	Contaminants possibles à ce type de sélection . . . . .	89
3.4	Adaptation de distribution spectrale d'énergie . . . . .	91
3.4.1	Historique et principe . . . . .	92
3.4.2	Comparaison aux diagrammes couleur-couleur . . . . .	96
<b>4</b>	<b>Méthodes de comparaison à des distributions spectrales d'énergie</b>	<b>99</b>
4.1	HyperZ - utilisation et qualification de l'outil. . . . .	100
4.1.1	Description du fonctionnement d'HyperZ . . . . .	100
4.1.2	Templates . . . . .	101
4.1.3	'Make catalogue' . . . . .	101
4.2	Tests et sélection des candidats . . . . .	102
4.2.1	Recherche de quasars . . . . .	102
4.2.2	Tests sur le SDSS . . . . .	104
	HyperZ et SDSS . . . . .	106
	Non détections . . . . .	108
	Absorption Lyman- $\alpha$ . . . . .	110
	Pas du décalage spectral . . . . .	112
	Choix des templates . . . . .	113
	Décalage spectral pour les tests . . . . .	118
	Critères de sélection . . . . .	119
	Nature des candidats . . . . .	123
4.2.3	Sélection sur le CFHTLS . . . . .	126
	Catalogues photométriques . . . . .	127
	Sélection . . . . .	128
	Suivi spectroscopique . . . . .	129
4.2.4	Conclusion . . . . .	130
<b>5</b>	<b>Milieu intergalactique</b>	<b>133</b>
5.1	Observations . . . . .	134
5.1.1	Forêt Lyman- $\alpha$ en observation . . . . .	135
5.1.2	Regroupement des raies . . . . .	136
5.2	Un milieu photo-ionisé . . . . .	136
5.3	Modèle actuel . . . . .	137
5.3.1	La forêt Lyman- $\alpha$ . . . . .	138
5.3.2	L'enrichissement en métaux du milieu intergalactique . . .	140
5.4	Tomographie du milieu . . . . .	140
5.4.1	Utilisation de paires de quasars . . . . .	141
5.4.2	Futur de la tomographie - les groupes de quasars . . . . .	141

5.5	Transverse correlation in the Lyman- $\alpha$ forest . . . . .	145
5.5.1	Introduction . . . . .	146
5.5.2	Observations . . . . .	147
5.5.3	Numerical simulations . . . . .	150
5.5.4	The observed correlation functions . . . . .	151
	Calculating correlation functions . . . . .	151
	The observed longitudinal correlation function . . . . .	152
	The observed transverse correlation function . . . . .	152
	Estimation of errors . . . . .	153
5.5.5	Comparison of observations and simulations . . . . .	154
	Systematic uncertainties . . . . .	154
	Observed <i>vs</i> simulated correlation functions . . . . .	156
5.5.6	Metal absorption systems . . . . .	158
	Identifying metal lines . . . . .	158
	The correlation of C IV systems . . . . .	160
	Peculiarities . . . . .	163
5.5.7	Conclusions . . . . .	164
<b>Conclusion</b>		<b>167</b>
<b>A Résumé</b>		<b>171</b>
<b>B Abstract</b>		<b>173</b>
<b>C Annexes de l'article</b>		<b>175</b>
C.1	Comments on individual lines of sight . . . . .	175
C.1.1	J 000852.7-290044–J 000857.7-290126 . . . . .	175
C.1.2	Q 0103-294A–Q 0103-294B . . . . .	175
C.1.3	Q 0236-2411–Q 0236-2413 . . . . .	175
C.1.4	J 023836.9-282310–J 023849.0-282101 . . . . .	175
C.1.5	UM 680–UM 681 . . . . .	176
C.1.6	J 031036.4-305108–J 031041.0-305027 . . . . .	176
C.1.7	FOCAP QSF :01–FOCAP QSF :04 . . . . .	176
C.1.8	J 095800.2-002858–J 095810.9-002733 . . . . .	176
C.1.9	J 102827.1-013641–J 102832.6-013448 . . . . .	176
C.1.10	Q 1206-1056–Q 1207-1057 . . . . .	176
C.1.11	J 120725.9-024519–J 120734.5-024725 . . . . .	177
C.1.12	J 123510.5-010746–J 123511.0-010830 . . . . .	177
C.1.13	J 125556.9+001848–J 125606.3+001728 . . . . .	177
C.1.14	J 135001.7-011703–J 135003.0-011819 . . . . .	177
C.1.15	J 141124.6-022943–J 141117.3-023222 . . . . .	177
C.1.16	J144245.6-024251–J144245.7-023906 . . . . .	177
C.1.17	Q 2129-4653A–Q 2129-4653B . . . . .	177
C.1.18	Q 2139-4504B–Q 2139-4504A . . . . .	178
C.1.19	J 214501.6-303121–J 214507.0-303046 . . . . .	178
C.1.20	J 223850.1-295612–J 223850.9-295301 . . . . .	178
C.1.21	J 232800.7-271655–J 232804.4-271713 . . . . .	178



C.2	Line Lists . . . . .	179
C.3	Spectra . . . . .	186
<b>D</b>	<b>HyperZ, ce qui a changé</b>	<b>193</b>
<b>E</b>	<b>'The' selection, how to..</b>	<b>197</b>
E.1	Photometric catalogues . . . . .	197
E.2	How to use HyperZ . . . . .	200
E.3	Final selection . . . . .	200
<b>F</b>	<b>List of templates</b>	<b>209</b>
	<b>Bibliographie</b>	<b>211</b>

# Table des figures

1.1	Spectre du quasar Q0453-423 . . . . .	23
1.2	Spectre de 3C273 . . . . .	26
1.3	Interprétation du spectre de 3C2732 . . . . .	26
1.4	Description des noyaux actifs de galaxie . . . . .	27
1.5	Émission radio des quasars . . . . .	28
1.6	Source à hautes énergies . . . . .	30
1.7	Diagramme couleur-couleur des quasars du 2QZ/6QZ . . . . .	34
2.1	Couverture photométrique du SDSS . . . . .	39
2.2	Comparaison magnitudes SDSS / logarithmique . . . . .	41
2.3	Base de données du SDSS DR1 . . . . .	43
2.4	Champs du CFHTLS . . . . .	47
2.5	Traitement des données du CFHTLS . . . . .	49
2.6	Caractéristique du relevé T0001 . . . . .	50
2.7	Caractéristique du relevé T0002 . . . . .	51
2.8	Filtres de MegaCam pour le CFHTLS . . . . .	53
2.9	Comparaison entre simulation et CFHTLS . . . . .	54
2.10	Comparaison entre simulation et CFHTLS - Térapix . . . . .	55
2.11	Comparaison entre simulation et SDSS . . . . .	56
2.12	Comparaison entre SDSS et CFHTLS . . . . .	57
2.13	Équation de couleur u-g SDSS et CFHT . . . . .	58
3.1	Spectre moyen sur les domaines d'énergie . . . . .	61
3.2	Comparaison radio quiet - radio loud . . . . .	62
3.3	Composite de quasar dans l'infra-rouge . . . . .	64
3.4	Composite optique / ultra-violet . . . . .	65
3.5	Exemple de systèmes métalliques - Q2139-4504B . . . . .	67
3.6	Exemple de BAL . . . . .	68
3.7	Introduction aux couleurs . . . . .	70
3.8	Spectres d'objets à travers les filtres SDSS . . . . .	71
3.9	Diagramme couleur-couleur pour le 2dF . . . . .	71
3.10	Simulation du locus des étoiles . . . . .	74
3.11	Simulation des couleurs des galaxies . . . . .	76
3.12	Simulation des couleurs des quasars . . . . .	77
3.13	Locus des quasars . . . . .	78
3.14	Algorithme de sélection du SDSS . . . . .	80
3.15	Le locus des étoiles dans le SDSS . . . . .	82

3.16	Régions d'exclusions pour la sélection de quasars . . . . .	83
3.17	Sélection simplifiée du SDSS . . . . .	88
3.18	Simulation du locus des étoiles - naines blanches . . . . .	90
3.19	Régions dans les diagrammes couleur-couleur . . . . .	91
3.20	Combinaison de filtre - HyperZ . . . . .	93
3.21	Adaptation dans COMBO-17 . . . . .	95
4.1	Distribution des quasars en fonction du décalage spectral dans le SDSS. . . . .	103
4.2	Densité de quasars dans les relevés . . . . .	105
4.3	Influence des erreurs sur HyperZ . . . . .	107
4.4	Influence des paramètres de filtres sur HyperZ - 1 . . . . .	109
4.5	Influence des paramètres de filtres sur HyperZ - 2 . . . . .	110
4.6	Traitement de la forêt Lyman- $\alpha$ par HyperZ . . . . .	111
4.7	Choix du pas de décalage spectral . . . . .	112
4.8	Template de quasar selon HST et SDSS . . . . .	114
4.9	Les templates des étoiles - 1 . . . . .	116
4.10	Les templates des étoiles - 2 . . . . .	116
4.11	Les templates des étoiles, exemple de spectres . . . . .	117
4.12	Efficacité des décalages spectraux photométriques . . . . .	118
4.13	Contaminants en fonction des décalages spectraux photométriques	119
4.14	Histogramme des candidats quasars en fonction du décalage pho- tométrique . . . . .	122
4.15	Candidats quasars : les fausses cibles . . . . .	124
4.16	Systèmes lorentzien : Diagramme couleur-couleur . . . . .	126
4.17	Quasars non identifiés par HyperZ : leurs couleurs . . . . .	127
4.18	Évolution du rayon des objets en fonction de leur magnitude . . .	129
4.19	Couleurs des quasars en fonction de leur décalage spectral . . . .	130
5.1	Exemple de forêt Lyman- $\alpha$ . . . . .	135
5.2	Simulation du milieu intergalactique . . . . .	139
5.3	Reconstruction du champ de densité . . . . .	143
5.4	Qualité de la reconstruction . . . . .	144
5.5	Typical observed spectra of QSO pairs . . . . .	148
5.6	The observed longitudinal correlation function - 1 . . . . .	153
5.7	The observed transverse correlation coefficient - 1 . . . . .	154
5.8	The longitudinal and transverse correlation functions . . . . .	156
5.9	The longitudinal correlation function - 2 . . . . .	157
5.10	The observed transverse correlation coefficient - 2 . . . . .	158
5.11	Histogram of the velocity separations-C IV systems . . . . .	160
5.12	Observed Longitudinal correlation function of C IV systems . . .	161
5.13	Spectra of the group . . . . .	162
C.1	Spectra of the observed QSO pairs - 1 . . . . .	186
C.2	Spectra of the observed QSO pairs - 2 . . . . .	187
C.3	Spectra of the observed QSO pairs - 3 . . . . .	188
C.4	Spectra of the observed QSO pairs - 4 . . . . .	189

C.5	Spectra of the observed QSO pairs - 5	190
C.6	Spectra of the observed QSO pairs - 6	191
E.1	belongpol.c - 1	202
E.2	belongpol.c - 2	203
E.3	hyperz_galaxies.param	204
E.4	hyperz_stars.param	205
E.5	hyperz_qsos.param	206
E.6	The different selections available	207



# Liste des tableaux

2.1	Paramètre $b$ et zéro flux magnitude pour le SDSS . . . . .	42
2.2	Champs observés au CFHT . . . . .	57
3.1	Principales raies d'émission dans l'optique et l'ultra-violet dans les spectres des quasars . . . . .	66
3.2	Combinaisons des filtres dans Hatziminaoglou et al. (2000) . . . .	93
3.3	Comparaison de DSE - Résultat de l'étude de Hatziminaoglou et al. (2000). . . . .	94
4.1	Densité de quasars . . . . .	104
4.2	Densité de quasars à faibles magnitudes . . . . .	104
4.3	Les différentes sélections - efficacité . . . . .	123
4.4	Sélection dans les champs profonds du CFHTLS . . . . .	131
5.1	Properties of the 33 QSO pairs . . . . .	159
5.2	Damped Lyman- $\alpha$ system candidates detected in the survey . .	165
E.1	content of the filter parameter . . . . .	200



# Préface

*"As far as I'm concerned, if something is so complicated that you can't explain it in 10 seconds, then it's probably not worth knowing anyway."*  
– Calvin

L'avènement des télescopes de la classe des 8-10m a permis d'obtenir des avancées majeures dans plusieurs domaines de la cosmologie aussi variés que l'étude de la formation des galaxies, leur évolution ou l'étude des systèmes d'absorption dans le spectre de quasars qui sont désormais à très haute résolution et à très grand rapport signal sur bruit. De plus, l'avancée de l'ère numérique a autorisé le stockage de grandes quantités de données dont l'accès est de plus en plus aisé, permettant une profusion d'études et de contre-études statistiques. Enfin les moyens de calcul accrus ont permis de construire des simulations de taille de plus en plus importante, allant jusqu'à réaliser des simulations hydro-dynamiques en trois dimensions, permettant une confrontation entre les modèles théoriques et les observations. Ainsi il n'est plus temps de considérer les domaines de la cosmologie moderne comme des nuages confinés par la pression de trop maigres observations ou par l'attraction trop forte de modèles invérifiables, mais comme de grands îlots reliés par les interactions de leurs objets de préoccupation et formant un tout indissociable convergeant vers **un** modèle et **une** histoire de l'Univers.

Las des grandes périphrases, il nous faut désormais attaquer la substance du texte et présenter à quoi il est destiné. Tout d'abord le contexte de cette thèse, celle-ci s'insère dans le projet de reconstruction en trois dimensions du milieu intergalactique, en d'autres mots, de réaliser la tomographie du milieu intergalactique afin de cartographier l'Univers. Pour se faire une idée de la naissance de ce projet, le lecteur est renvoyé aux textes de Patrick Petitjean (1997, 1999) ; pour une application de la reconstruction du milieu intergalactique il est conseillé de se reporter à Pichon et al. (2001).

La tomographie du milieu intergalactique est un projet qui nécessite une certaine densité de quasars ; Nusser & Haehnelt (1998) suggèrent d'utiliser le relevé proposé par Petitjean (1997), qui consiste en une collecte de spectres de résolution moyenne jusqu'à une magnitude de 22 dans un champ de deux degrés carrés. Un tel relevé permettrait d'obtenir à peu près une centaine de lignes de visée vers des quasars dans une région d'une trentaine de méga parsec de section. Mais pour arriver à cet objectif, il faut se concentrer sur des sur-densités de quasars, difficiles à trouver et pas toujours représentatives du milieu intergalactique. L'autre solution consiste à aller à des magnitudes plus profondes, et donc



utiliser de nouveaux relevés. En attendant que l'on collecte suffisamment de ces lignes de visées groupées, il est possible de mener des études de reconstruction à partir de paires de quasars, qui sont rares mais tout de même plus fréquentes que les groupes.

C'est ainsi que va se découper ce mémoire, une partie présentant la recherche des quasars afin d'aller à des magnitudes de plus en plus faibles pour obtenir des densités de plus en plus grandes et une partie présentant les outils d'analyse du milieu intergalactique à l'aide des paires de quasars.

L'introduction présentera le contexte général dans lequel se situe cette étude, du point de vue de nos connaissances sur le milieu intergalactique, sur les bases de données des quasars et sur leur recherche.

Le principe même de l'identification d'objets, les quasars pour ce qui nous concerne, est de comparer une base de connaissances à de nouvelles observations. Le chapitre 2 montrera quels sont les relevés qui vont nous permettre une avancée importante dans la collecte des lignes de visée. Les bases de connaissances sont l'ensemble des spectres de quasars collectés jusqu'à aujourd'hui (pour l'essentiel provenant du SDSS, Sloan Digital Sky Survey) et les nouvelles observations qui sont utilisées dans cette thèse proviennent du relevé en imagerie CFHTLS (Canada France Hawaï Telescope Legacy Survey).

Les méthodes de sélection des quasars ont beaucoup évolué ces dernières décennies et seront présentées dans le chapitre 3. Dans ce dernier nous verrons que la méthode d'adaptation de distribution spectrale d'énergie (DSE) semble plus adaptable et plus fiable d'un relevé à un autre; elle se montre moins spécifique au relevé en présentant des résultats comparables.

Les tests menés sur le SDSS avec HyperZ présentés dans le chapitre 4 permettent d'estimer dans un premier temps l'efficacité de notre sélection. Les tests effectués sur les bases du SDSS nous aident à disposer de plusieurs types de sélection selon les stratégies des suivis spectroscopiques. Pour chacune d'entre elles nous avons une estimation de l'efficacité. L'auteur présente comment ces critères de sélection peuvent être appliqués au relevé CFHTLS et quelles en sont les contraintes. Le dernier chapitre représente la base du travail qu'il faudra effectuer une fois que l'on aura collecté suffisamment de lignes de visée. En effet nous avons montré qu'avec un nombre faible de paires de quasars, il était possible d'accorder observations et simulations, ainsi que d'obtenir un échantillon suffisant pour l'estimation de la corrélation longitudinale.

# Chapitre 1

## Introduction

*"The only skills I have the patience to learn are those that have no real application in life."*  
– Calvin

## Sommaire

---

<b>1.1</b>	<b>Le milieu intergalactique, son histoire, son étude à travers les spectres des quasars . . . . .</b>	<b>20</b>
<b>1.2</b>	<b>Comment rechercher les quasars ? . . . . .</b>	<b>24</b>
1.2.1	Cette belle histoire qu'est la recherche de quasars . . .	25
	Description des quasars. . . . .	25
	Les catalogues existants de quasars . . . . .	27
1.2.2	Méthodes de recherche suivant les différents domaines d'énergie . . . . .	32
	Recherche dans le domaine optique . . . . .	33
	Recherches venues d'ailleurs . . . . .	35

---

Cela fait désormais plus de trois quarts de siècle que nous avons conscience que notre Univers ne se résume pas à notre galaxie et qu'il est en expansion. Cette prise de conscience est survenue alors que les premiers fondements de la cosmologie actuelle étaient énoncés, notamment le principe cosmologique postulant que l'univers est homogène et isotrope pourvu que l'on considère une échelle suffisamment grande. A partir de celui-ci il a été possible à Alexander Friedmann d'écrire les équations (provenant de la théorie de la relativité générale) gouvernant l'évolution de l'univers dans un modèle de 'Big-Bang'.

Aujourd'hui les observations nous permettent de remarquer l'inhomogénéité de l'univers sur certaines échelles ; et il nous reste à comprendre comment l'univers a pu passer d'un état quasi uniforme lors de sa formation, à l'aspect complexe que nous lui connaissons maintenant, où la matière s'est agrégée jusqu'à former les galaxies et les structures encore plus petites. Il convient donc pour répondre à cette question de jouer aux historiens de l'univers, de partir à la recherche des preuves qu'il est encore possible de récolter, à partir de l'émission du fond diffus cosmique (relique d'émission qui provient d'une époque où l'univers était presque totalement homogène et en équilibre thermique) et de déterminer le moment où les premiers objets se sont formés via la gravitation. Le temps historique se transforme pour les chercheurs en temps 'cosmologique' et celui-ci se mesure en décalage spectral, noté  $z$ . La quantité  $1+z$  est le facteur par lequel l'univers s'est étendu entre cette époque mesurée et la nôtre. Cette quantité correspond aussi au facteur par lequel la longueur d'onde de la lumière émise à cette époque par n'importe quel objet, et nous parvenant aujourd'hui, a été étirée, dû à l'expansion de l'univers.

## 1.1 Le milieu intergalactique, son histoire, son étude à travers les spectres des quasars

A un très grand décalage spectral (à très grand  $z$ ), l'univers était presque complètement homogène, la température de la matière diminuait en même temps que l'univers s'étendait. Le plasma se recombina, c'est-à-dire que les atomes se formèrent à  $z \simeq 1100$  lorsque la température était de  $T = 3000K$  : cette période

est appelée la recombinaison. C'est à cette époque que le fond diffus cosmologique remplit l'univers d'une radiation dont le spectre était celui d'un corps noir à cette température. Plus tard, la température décrut et le spectre du fond diffus cosmologique se trouva décalé vers le rouge. Après ce découplage rayonnement-matière et l'émission de cette radiation, la matière était dans un état complètement neutre. En effet, aucun observateur présent sur place (à cette époque) n'aurait pu voir d'éventuels photons émis, l'univers étant alors complètement neutre, tout photon émis était immédiatement et irrémédiablement absorbé. Pour éclairer cet univers complètement noir, il a fallu attendre longtemps que les premiers objets se forment dans ce milieu neutre et froid, sous l'effet de la gravitation. En effet, durant cette période noire (parfois appelé 'l'âge sombre', Barkana & Loeb 2001) le gaz atomique était toujours quasi-homogène et seulement une très faible fraction de ce gaz a pu se transformer en molécules telles que  $H_2$ ,  $HD$  et  $LiH$  lorsque la température diminua et où la densité était suffisante.

Les observations effectuées dans le cadre de la cosmologie actuelle renforcent le paradigme des modèles de matière noire froide. Ces modèles supposent qu'en plus de la matière ordinaire, faite de protons, neutrons et électrons, c'est-à-dire la matière baryonique, il existe aussi une matière dite 'noire'. Celle-ci serait constituée de particules ne subissant aucune autre interaction que celle de la gravité et qui étaient initialement froides (i.e. les particules avaient une dispersion en vitesse très faible). Cette matière noire froide est nécessaire pour expliquer les fluctuations primordiales de faible amplitude qui ont ensuite grandi jusqu'à former les plus grandes structures (galaxies, amas...). Ceci nous amène à la formation des premières étoiles. En effet, l'âge sombre s'est terminé avec la formation de celles-ci. Pour pouvoir former ces étoiles, le gaz atomique doit suivre, à cause de la gravitation, la matière noire qui a déjà formé des halos. Les effondrements ne sont possibles que dans les puits de potentiel suffisamment grands. Ceci arrive lorsque la masse du halo est plus grande que la masse de Jeans du gaz <sup>1</sup> à la température virialisée <sup>2</sup> et pour la densité du milieu intergalactique, ce qui est satisfait lorsque  $T_{vir} \geq 100K$  (Blumenthal et al. 1984, Couchman et al. 1986). Dans les halos de température plus basse, la pression du gaz est suffisante pour empêcher le gaz de s'effondrer. De plus il faut que le gaz se refroidisse jusqu'à ce que des densités suffisantes soient atteintes pour former des étoiles. Ce refroidissement dépend de  $T_{vir}$  du halo et de la composition du gaz. Le gaz primordial était essentiellement composé d'hydrogène et d'hélium atomiques. L'hydrogène atomique ne peut refroidir le gaz que lorsque  $T_{vir} > 10^4 K$ , c'est-à-dire lorsque les collisions permettent d'exciter et de ioniser les atomes d'hydrogène <sup>3</sup>, le gaz peut alors se contracter et former les premiers objets. Si l'on considère un intervalle intermédiaire de  $T_{vir}$ , par exemple  $100K < T_{vir} < 10^4 K$ , le refroidissement atomique du gaz ne peut pas se réaliser,

---

<sup>1</sup>La masse de Jeans d'un gaz à une certaine température et densité est la masse minimale requise pour que la gravité dépasse le gradient de pression et amène ainsi le gaz à s'effondrer.

<sup>2</sup> $T_{vir}$  qui est proportionnelle à l'énergie cinétique du gaz.

<sup>3</sup>Les atomes d'hélium nécessitent encore plus d'énergie pour réaliser une transition jusqu'au premier niveau d'excitation que les atomes d'hydrogène, et donc l'hélium ne permet un refroidissement qu'à des températures plus élevées que pour l'hydrogène.

et en l'absence des éléments lourds qui seront éjectés plus tard par les étoiles massives, le seul élément permettant le refroidissement est  $H_2$ . Et seule une très faible partie du gaz dans ces premières formations a pu se transformer en  $H_2$  (Saslaw et al. 1967, Tegmark et al. 1997) ce qui limite le taux de refroidissement du gaz. D'après les simulations, les premières étoiles se seraient formées dans ces conditions dans des halos de température  $T_{vir} \simeq 2000K$  et auraient eu une masse de  $\sim 10^6 M_\odot$ . Ces premières étoiles modifièrent leur environnement en ionisant celui-ci et en l'enrichissant en métaux, notamment par les toutes premières supernovae. Celles-ci ont aussi pu accélérer le gaz et le faire tomber dans des objets encore plus massifs. La question qui se pose alors est : quand se sont formées les premières étoiles ? La réponse n'est pas simple, et requiert l'analyse des fluctuations de densité primordiale. Pour cela l'étude complète du spectre de puissance est nécessaire afin de déterminer leur fréquence et leur amplitude. Selon ces facteurs on estime que les premières étoiles ont pu se former entre  $z \simeq 30$  et  $z \simeq 40$  dans les modèles de matière noire froide.

L'effet le plus important qu'a eu la formation d'étoiles (ou d'autres objets plus massifs) sur leur environnement fut la réionisation du gaz dans l'univers. Lorsque les premiers objets se formèrent, les photons émis furent immédiatement absorbés par le gaz à très haute densité résidant dans le halo. Ceci est dû au fait que les photons ionisants dans l'ultraviolet ont un parcours moyen très court à travers ce gaz essentiellement atomique. Le milieu intergalactique commença à être réionisé lorsque des sources suffisamment puissantes purent ioniser tout le gaz dans leur propre halo, permettant aux photons ionisants de s'échapper. La réionisation s'est donc déroulée par l'expansion de 'bulles' de réionisation autour des sources, séparant l'univers en 'bulles' ionisées et un milieu atomique entre les bulles (Arons et al. 1972). Ces bulles ionisées ont grandi et se sont rejointes jusqu'à ce que les régions les moins denses de l'univers soient ionisées. Au final le milieu est principalement ionisé et on peut encore observer des résidus neutres en absorption dans les spectres d'objets lointains.

Les preuves de cette réionisation proviennent de l'observation des spectres de quasars. Je reviendrai plus en détail sur ce que sont les quasars ; en résumé ce sont des objets extrêmement lumineux qui se trouvent au centre de galaxies lointaines et qui sont alimentés par l'accrétion de matière autour de trous noirs super massifs. Grâce à leur très grande luminosité, ils sont utilisés par les cosmologistes comme des lampes dont les spectres permettent, via l'étude des raies d'absorption, d'extraire des informations sur l'état de la matière située le long de la ligne de visée. Les spectres des quasars montrent la présence de lumière en-deçà de la longueur d'onde d'émission Lyman- $\alpha$  de l'hydrogène. Or, si le milieu intergalactique était complètement sous forme d'atomes et homogène, tout photon émis à une longueur d'onde plus courte que Lyman- $\alpha$  (121.6 nm) serait absorbé. Cette suppression du flux à plus petite longueur d'onde que l'émission Lyman- $\alpha$  est appelée l'effet Gunn-Peterson (1965). Dans les quasars situés à un décalage spectral inférieur à 6, cet effet n'est pas retrouvé. A la place on observe un flux partiellement absorbé, ce qui correspond à ce qui est appelé la forêt Lyman- $\alpha$  : un grand nombre de raies d'absorptions de forces différentes le long du spectre. Les atomes d'hydrogène dans le milieu intergalactique produisant ces absorptions représentent une petite fraction de l'hydrogène total, qui

est en équilibre de photo-ionisation avec le fond cosmique ionisant produit par les galaxies et les quasars (Rauch 1998). Les raies d'absorption correspondent aux variations de densité de la matière dans le milieu intergalactique. Ces observations montrent que la majorité des baryons se trouvent sous forme ionisés. De plus, l'étude des spectres de quasar a révélé que la majorité des baryons dans l'univers, à un décalage spectral supérieur à 3, se trouve dans le milieu intergalactique.

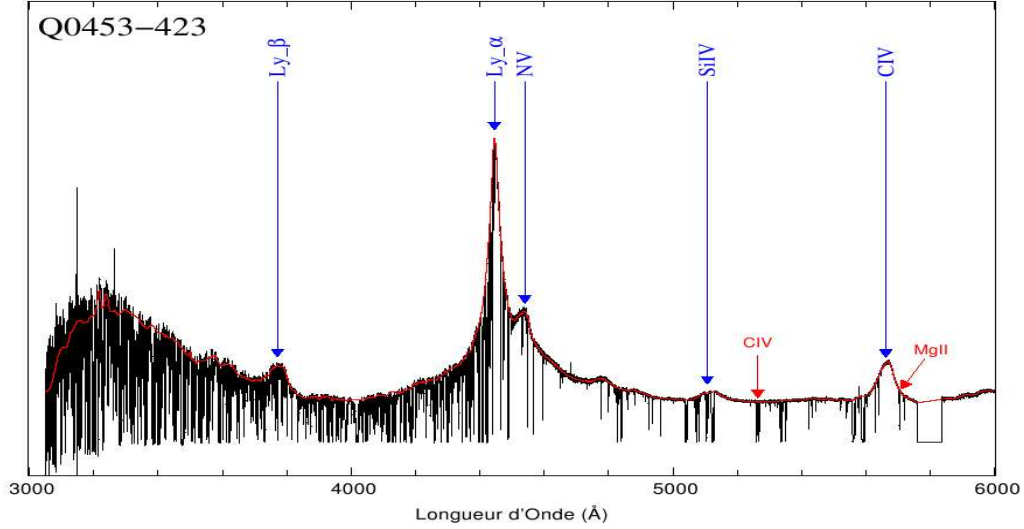


FIG. 1.1 – Spectre du quasar Q0453-423 observé avec UVES. Figure issue de la thèse de Bastien Aracil.

Un spectre de quasar en émission est essentiellement constitué de raies d'émission très fortes et larges, un continu en loi de puissance dans les domaines radio, X, ultra-violet, optique et infra-rouge (voir partie 3.1). Le spectre en absorption qui se superpose au continu est dû au milieu intergalactique qui laisse son empreinte dans le spectre des quasars observés, notamment l'hydrogène neutre via la forêt Lyman- $\alpha$ . Les métaux qui composent en moindre partie le milieu intergalactique nous renseignent essentiellement sur la proximité de sur-densités ou d'objets tels que des galaxies à proximité de la ligne de visée du quasar.

Je vais passer en revue rapidement quelques propriétés générales des absorbants dans les spectres des quasars, à quoi ils sont reliés et comment ils peuvent nous apporter des informations sur le milieu intergalactique. Les absorbants qui peuvent être vus dans les spectres de quasars doivent forcément être composés d'éléments ayant une transition atomique qui permet de les observer avec les moyens technologiques aujourd'hui disponibles. Dans certains cas il est possible d'observer l'absorption par des molécules. Comme le milieu est dans sa plus grande partie ionisé, il est essentiellement transparent sauf dans l'ultra-violet où l'on peut voir la transition Lyman- $\alpha$  en absorption. Il est possible de détecter des raies situées dans l'ultra-violet au moment de l'absorption car ce que l'on observe est en fait la même transition décalée d'un facteur  $(1+z)$ ,  $\lambda_{obs} = (1+z)\lambda_{lab}$ . Typiquement pour la transition Lyman- $\alpha$ , l'observation peut se faire à partir

d'un décalage spectral ( $z$ ) de 1.8 approximativement.

Les absorbants peuvent généralement être classifiés selon leur densité de colonne qui correspond au nombre d'atomes par unité de surface présents le long de la ligne de visée. La forêt Lyman- $\alpha$  étant due essentiellement au milieu intergalactique, les absorbants sont vus dans la partie bleue de l'émission Lyman- $\alpha$  du quasar, et ont typiquement une densité de colonne  $N(\text{HI})$  de l'ordre de  $10^{14}$  atomes/cm<sup>2</sup>.

Mais la diversité des raies observées est grande, des faibles densités de colonne jusqu'à des densités supérieures à  $10^{20}$  atomes par centimètre carré, on parle dans ce cas de systèmes lorentziens. Ces derniers ont des densités de colonne semblables à celles des galaxies spirales d'aujourd'hui ou à leurs progéniteurs. De même les raies les plus fortes de la forêt, comprenant généralement des métaux tels que C IV ou Mg II sont associées à des halos de galaxies. A partir de l'observation des absorptions dans les spectres des quasars, il est possible de récolter des informations sur la densité du gaz formant le milieu intergalactique et de retracer la formation des structures.

De plus les simulations nous apportent un moyen efficace de comparer les observations aux théories en traitant gaz et matière noire simultanément. L'une des études les plus importantes menées sur le milieu intergalactique fut la recherche de la taille des absorbants. Pendant longtemps on avait en tête l'image d'absorbants comme étant des nuages, confinés par la gravitation, pression, etc... Afin de contraindre la taille de ces nuages il a fallu utiliser des paires de quasars (dans le sens paire projetée sur le ciel). Car ainsi on obtenait non pas une information uniquement le long de la ligne de visée mais aussi dans la direction transverse. Les recherches menées par Smette et al. (1992), Fang et al. (1996) montrent que de tels nuages ont des tailles de l'ordre du méga parsec. Avec des volumes aussi grands, la densité, elle, est très faible ce qui implique un fort degré d'ionisation et donc la partie du gaz que l'on observe est une partie infime du gaz présent. Désormais les modèles du milieu intergalactique montrent qu'il ne s'agit pas de nuages mais de filaments de gaz. Mais l'idée d'utiliser une information transverse ne se restreint pas à contraindre la taille des absorbants. En comparant l'information transverse et longitudinale sur un intervalle en décalage spectral suffisamment grand, on peut appliquer le test d'Alcock et Paczynski (1979) pour déterminer les paramètres cosmologiques (McDonald 2003). Il est possible d'aller encore plus loin en utilisant des groupes de quasars, qui peuvent nous permettre de réaliser une véritable tomographie du milieu intergalactique (Pichon et al. 2001). Je décris tout ceci dans le chapitre 5.

Mais pour observer des groupes de quasars avec des séparations de l'ordre de 10 arcmin et moins, il faut se donner la peine de les chercher...

## 1.2 Comment rechercher les quasars ?

Si les phénomènes qui sont à la source du rayonnement du quasar déterminent en partie leur aspect, les absorptions influent sur la quantité de lumière qui nous parvient. Ainsi la reconnaissance des quasars et leur différenciation à partir de leur couleur et de leur luminosité dans une partie du spectre, dépend

à la fois de ce qui engendre leur rayonnement et du milieu traversé par le flux lumineux.

L'une des principales raies d'émission, la plus forte du spectre des quasars, est la raie d'émission Lyman- $\alpha$  de l'hydrogène neutre. Et selon le décalage spectral du quasar il est possible de détecter cette raie dans l'ultra-violet, le visible ou l'infra rouge. Ainsi il est possible de chercher des quasars en se fiant à leurs couleurs (qui sont très liées à la raie d'émission Lyman- $\alpha$  mais aussi aux autres raies et au continu). Ils peuvent par exemple paraître plus bleus que la plupart des étoiles si cette raie a été décalée vers le rouge jusqu'à atteindre des longueurs d'onde de l'ordre de 3800 Å. Lors d'un sondage du ciel dans différents filtres, il est possible de mesurer des excès de couleurs, reflétant la présence ou l'absence possible d'une forte raie d'émission dans un des filtres. Les quasars ont aussi la propriété d'avoir une luminosité qui varie (permettant de les distinguer d'objets à la luminosité constante), ainsi, on peut par exemple avec un sondage sur plusieurs époques rechercher les objets qui ont vu leur luminosité varier d'une observation à l'autre.

La sélection des quasars est donc liée à la fois à leur nature, à leur environnement direct mais aussi au milieu intergalactique qui influe sur leur couleur et leur éventuelle détection lors des sondages du ciel. **Le principe de base d'une sélection est de rechercher des objets d'après l'idée que l'on se fait d'eux a priori.** Il s'agit de les décrire d'une certaine façon, par des caractéristiques fortes, pour ensuite rechercher les objets dans le ciel qui correspondent à cette description. Il est donc nécessaire dans un premier temps de décrire le plus précisément possible et correctement les quasars.

### 1.2.1 Cette belle histoire qu'est la recherche de quasars

#### Description des quasars.

La découverte, ou plutôt l'identification des quasars (quasi-stellar object) date de Schmidt (1966).

Il s'agissait essentiellement d'expliquer les étonnantes raies d'émission observées dans le spectre. L'explication est venue de la coïncidence entre les raies observées et les raies de l'hydrogène, mais décalées vers le rouge, décalage qu'il a fallu alors justifier. Différentes théories ont vu le jour, comme une vitesse d'éjection très importante (théorie persistante à propos de l'éjection des quasars à partir d'une galaxie, e.g. Bell et al. 2004). Il est finalement apparu clairement que le décalage est dû à l'expansion de l'univers.

Désormais il existe un modèle décrivant les quasars, et d'une façon plus générale la classe des noyaux actifs de galaxies (NAG), dont font partie les quasars. La description schématique est donnée figure 1.4. Le gaz présent dans les régions denses autour du noyau (constitué d'un trou noir super massif et des régions internes du disque d'accrétion) émet les raies larges (broad line region, BLR) tandis que les régions plus éloignées sont responsables de l'émission de raies étroites (narrow line region, NLR), au décalage spectral du quasar. Ce schéma de la coupe d'un noyau actif de galaxie le long de l'axe logarithmique normalisé au parsec montre les constituants essentiels : le trou noir entouré



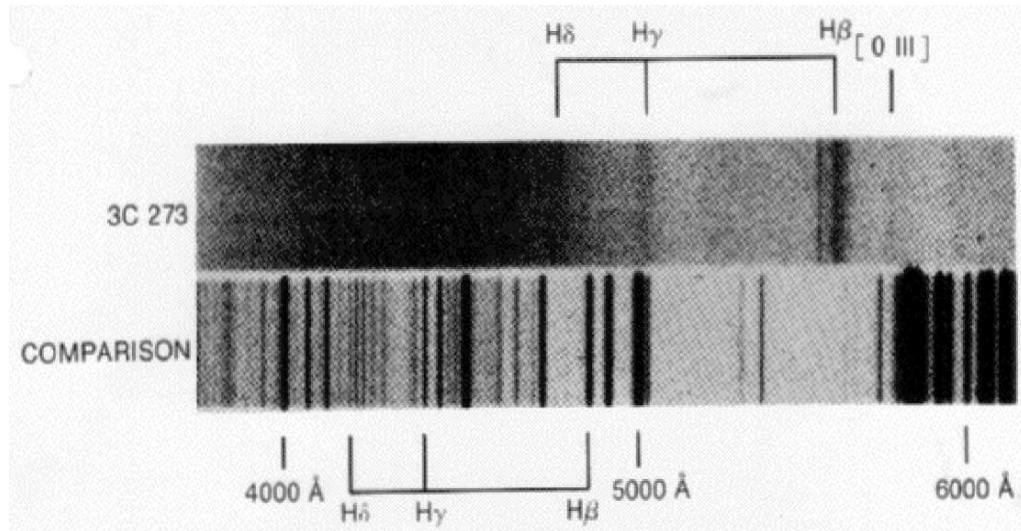


FIG. 1.2 – Spectre du quasar 3C273 tel qu'il a été publié par Schmidt.

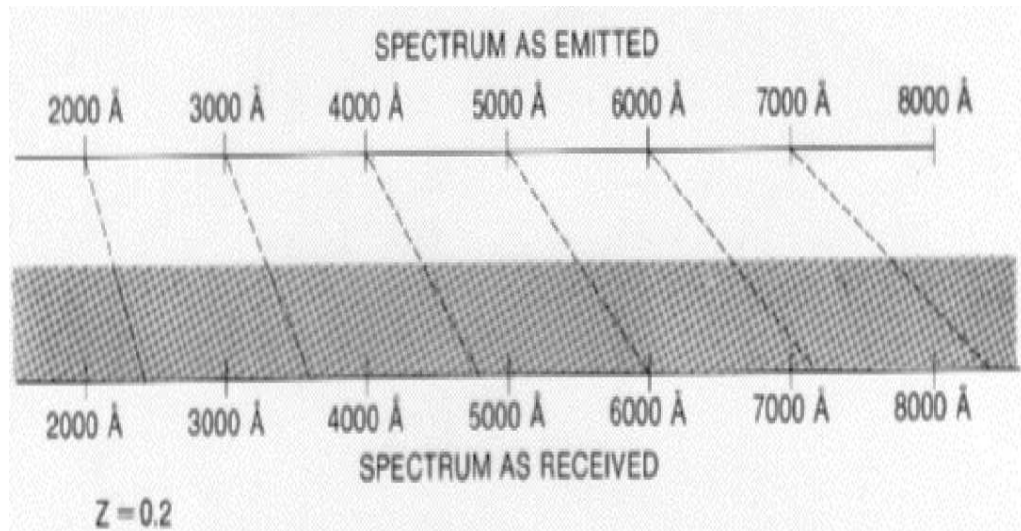


FIG. 1.3 – Interprétation des raies décalées du spectre du quasar 3C273. 3C273 a désormais un site bien à lui : <http://isdc.unige.ch/3c273/>.

du disque d'accrétion ainsi que la localisation de la BLR qui s'étend sur à peu près 1 pc et la NLR sur une centaine de pc. Le tout est entouré d'un tore de poussières qui selon l'angle de vue permet à l'observateur de voir en direction du noyau central ou pas. Selon cet angle de vue, la classe de noyau actif de galaxie change. L'observation typique d'un quasar en radio est présentée figure 1.5, où l'on peut voir les différentes régions d'émissions en radio, la zone centrale d'une part et parfois les lobes d'autre part qui sont liés à des jets de matière provenant de la zone d'accrétion.

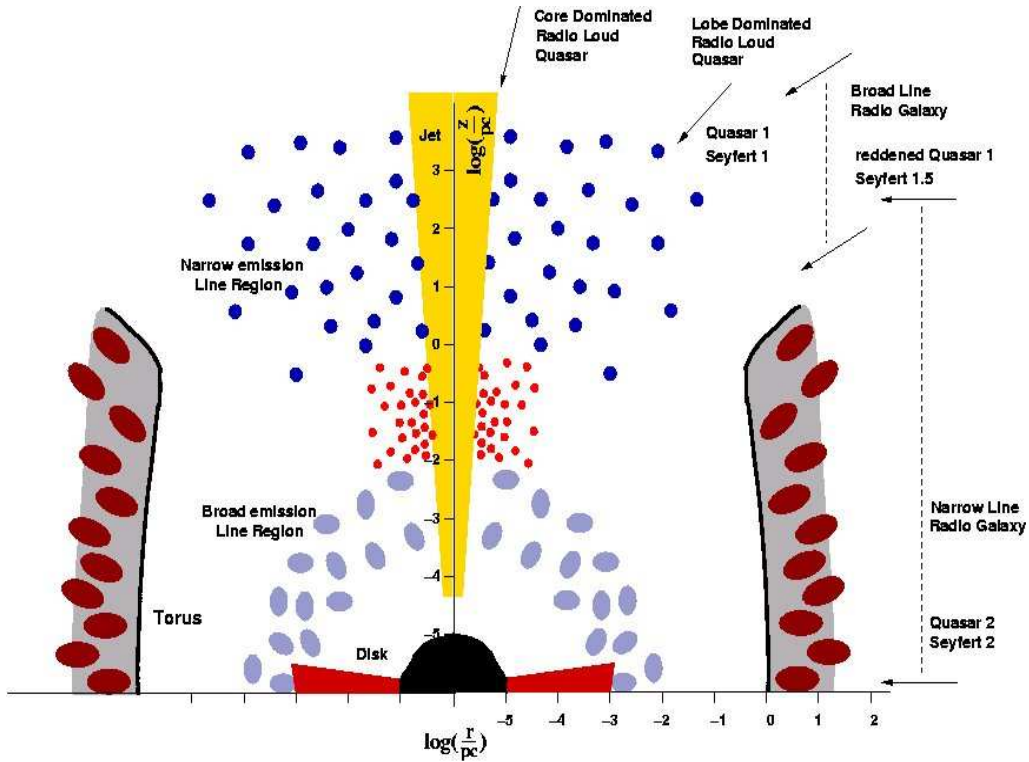


FIG. 1.4 – Représentation en échelle logarithmique des noyaux actifs de galaxies. Suivant l'angle de vue que l'on a sur cet objet, il appartient à différentes classes.

### Les catalogues existants de quasars

#### Catalogues issus de relevés en radio

Le premier catalogue de quasars réalisé le fut à partir de données radio, le Third Cambridge Catalogue (3C). Tout le catalogue ne contient pas, bien sûr, que des quasars, les sources ont dû être identifiées ensuite par des observations en optique et discerner les quasars des autres sources radio. Il est à noter pour la petite histoire qu'il n'y a pas eu de premier ou de deuxième catalogue. En tout cas pas de catalogue utilisable, les astronomes de l'époque ayant été confrontés aux problèmes de l'observation en radio. En fait un grand nombre de sources se sont révélées être des sources multiples non "résolues", et une détection correspondait en fait à plusieurs sources. Le troisième catalogue fut réalisé à une fréquence de 159 MHz, ce qui permit d'éliminer ce problème. Il comprend 471 sources entre -22 et +71 degrés, avec une densité de flux supérieure à 8 Jy. Le catalogue fut révisé par Bennett en 1962. La révision était requise car certaines sources devaient être retirées ou ajoutées car leur flux était sous la limite de 8 Jy ou furent identifiées comme des sources très proches qui n'étaient alors pas résolues. Ces catalogues furent ensuite suivis par le 4C (Pilkington & Scott (1965) et Gower, Scott & Wills (1967)), le 5C publié en plusieurs parties entre 1975 et 1995. (Pearson 1978 et Benn 1991), le 6C à 151-MHz publié entre 1985 et 1993.

Les relevés effectués par Cambridge valent pour l'hémisphère nord ; en ce qui concerne le sud, des relevés identiques ont été effectués aux mêmes dates par les

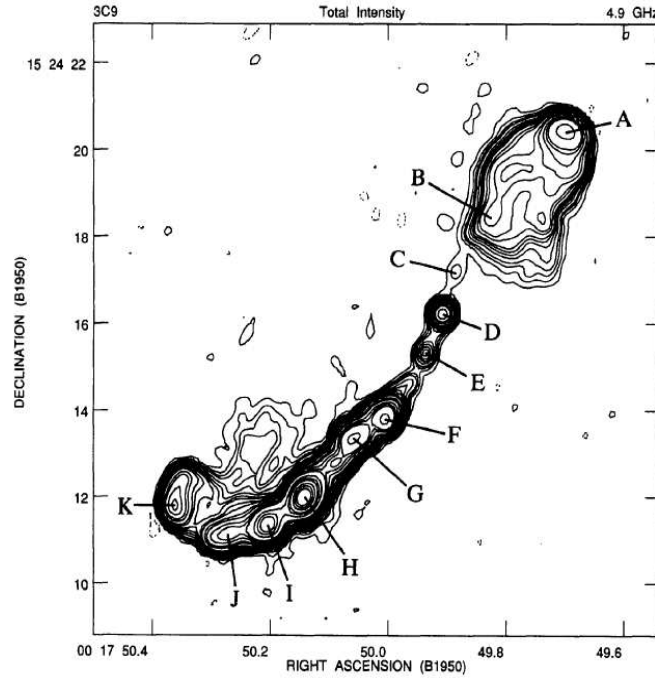


FIG. 1.5 – Émission radio du quasar 3C334 se situant à un décalage spectral de 0.555. L'image a été prise au VLA à 4.9 GHz avec 0.3 arsec de résolution. Le point (D) est le quasar, aussi identifié en optique. Issue de Bridle et al. 1994

australiens du Parkes radio telescope (c'est cette équipe qui mit en évidence les problèmes rencontrés par l'équipe de Cambridge). La fréquence d'observation était de 408 MHz et une sensibilité jusqu'à 3 Jy en limite de flux. Les relevés fait à ce radio télescope ont été compilés dans un catalogue, le PKSCAT90 (<http://www.parkes.atnf.csiro.au/research/surveys/pkscat90.html>).

Plus récemment les relevés tel que le NVSS (NRAO VLA Sky Survey) s'effectuent à 1.4 GHz couvrant la partie du ciel au nord de -40 degrés de déclinaison. (Condon et al. 1998). Le relevé FIRST (Faint Images of the Radio Sky at Twenty-cm) fut prévu pour être l'équivalent radio du Palomar Observatory Sky Survey sur 10 000 degrés carrés dans l'hémisphère nord. Il fut réalisé avec le NRAO Very Large Array (VLA), centré sur 1365 et 1435 MHz jusqu'à une limite de détection de 1 mJy, le catalogue contient 811 000 objets. 15% des sources ont une contrepartie dans les limites du relevé optique POSSI. Le relevé a été choisi pour coïncider avec le SDSS (Sloan Digital Sky Survey) qui sera décrit en détail (cf. 2.1) où on estime que 50% des sources de FIRST seront détectées.

### Catalogues issus de relevés en X

Le fond cosmique en rayon X a été étudié en 1962 (Giacconi et al.) lors du lancement d'une fusée. Ce ne fut pas un réel relevé, il s'agissait de trouver des sources de rayonnements X extra-solaires. Les premiers relevés ont été effectués dans les années soixante-dix, Uhuru et Ariel V. Ceux-ci ont montré que le fond X était essentiellement extragalactique. La nature de ce fond cosmique commença

à être élucidée grâce à l'apparition des télescopes de type Wolter qui améliorèrent considérablement la résolution angulaire et révélèrent la nature discrète du rayonnement X. Notamment que les NAG connus sont des émetteurs en rayon-X. (Tananbaum et al. 1979). En effet, les observations faites avec le satellite Einstein résolurent 25% des sources à 1-3 keV jusqu'à  $3.10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ , et furent identifiées comme NAG. Ceci fut confirmé par ROSAT (1993, 1998) à 0.5-2 keV (300 Å- 60 Å) jusqu'à des sensibilités de  $5.5.10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$  (Hasinger et al. 1998). En effet, ce dernier relevé montra une densité de NAG de l'ordre de 850 par degré carré. Pour des énergies plus grandes, il fallut attendre ASCA qui fut un relevé à 2-7 keV jusqu'à  $5.10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ . Puis le pionnier du domaine d'énergie 5-10 keV : BeppoSAX qui résolut 30% des sources à ces énergies (Comastri et al 2001.).

Les relevés d'aujourd'hui sont effectués par les satellites Chandra (NASA, Weisskopf et al. 2000) et XMM-Newton (ESA, X-ray Multi-Mirror Mission Newton, Jansen et al. 2001) équipés de télescopes Wolter. Ils disposent tous les deux de détecteurs pour atteindre les énergies 0.5-10 keV avec une sensibilité de 50 à 250 fois plus grande que les relevés précédents. Les différentes sources ont une bonne résolution astrométrique, 0.3-1" pour Chandra et 1-3" pour XMM-Newton, essentielle pour pouvoir corrélérer ces catalogues à ceux réalisés en optique (SDSS, CFHTLS etc...). Ces deux missions sont complémentaires en termes de résolution et de sensibilité. Dans le domaine des rayons X, les relevés avant Chandra et XMM ne disposaient pas d'une résolution ni d'une sensibilité suffisante pour se dédier à la recherche de NAG. Aux limites de ces surveys, les NAG de faible luminosité sont détectés, jusqu'à un décalage spectral de 4 approximativement. Les sources brillantes de ces relevés sont essentiellement des NAG non-obscuris et sont faciles à identifier par spectroscopie. Les NAG obscuris sont plus difficiles à identifier en optique/UV mais montrent des caractéristiques typiques en X. Les sources trop faibles pour être confirmées en optique ( $I > 24-25$ ) peuvent néanmoins être en grande partie des NAG. Viennent ensuite les galaxies normales, que l'on soupçonne d'avoir une faible activité de type NAG. Les prochains relevés se devront d'améliorer la sensibilité, d'ici 10-20 ans on attend XEUS (X-ray Evolving Universe Spectroscopy) et Generation-X.

### Catalogues issus de relevés en infra-rouge

En ce qui concerne les domaines d'énergie infra-rouge, les relevés comme IRAS ont permis la découverte de NAG, même si cela n'était pas le but premier du satellite. Et la recherche d'objets particuliers en infra-rouge a permis la découverte de nouveaux types de galaxies, les galaxies 'ultra-lumineuses' IRAS, dont le mécanisme d'émission pourrait inclure à la fois une formation d'étoiles importante et un noyau actif. A noter les relevés dans l'infrarouge proche, 2MASS (Skrutskie et al. 1997) et DENIS (Epchtein et al. 1997) qui complète les grands relevés optiques.

### Catalogues issus de relevés en rayon gamma

Explorer les domaines d'énergie encore plus grande se révèle compliqué. En effet, les observatoires de rayon gamma sont difficiles à construire, et il est

particulièrement difficile d'extraire un signal de la grande quantité de bruit qui accompagne les observations. De plus la résolution ne serait-ce que de un degré est difficile à atteindre. Mais le jeu en vaut la chandelle, même s'ils ne sont pas directement liés aux quasars, les objets trouvés sont toujours particuliers. Une des détections les plus marquantes a été réalisée lors de l'utilisation du télescope gamma pour déterminer la nature du centre galactique, qui permet en fait de découvrir les micro-quasars. Une image prise par INTEGRAL (cf. un peu plus loin) est présentée figure 1.6.

La première collection d'objet, à travers une recherche systématique d'émetteurs gamma fut menée par EGRET (Energetic Gamma-Ray Experiment Telescope, lancé en 1991 par la NASA) emportant le Compton gamma-ray observatory. Il couvrit tout le ciel, et permit de résoudre certains NAG jusqu'à  $10^{-11} \text{ erg cm}^{-2} \text{ s}^{-1}$  dans le domaine d'énergie 30 MeV-3GeV. Mais une fois encore la résolution est très faible, de l'ordre du degré dans le meilleur des cas. Aujourd'hui le principal satellite dans ce domaine est INTEGRAL (INTERnational Gamma-Ray Astrophysics Laboratory, lancé en 2002 par l'ESA). A noter les télescopes sur Terre tels que HESS (High Energy Stereoscopic System), MAGIC. Ainsi que le satellite SWIFT qui est consacré à la recherche de sursaut gamma.

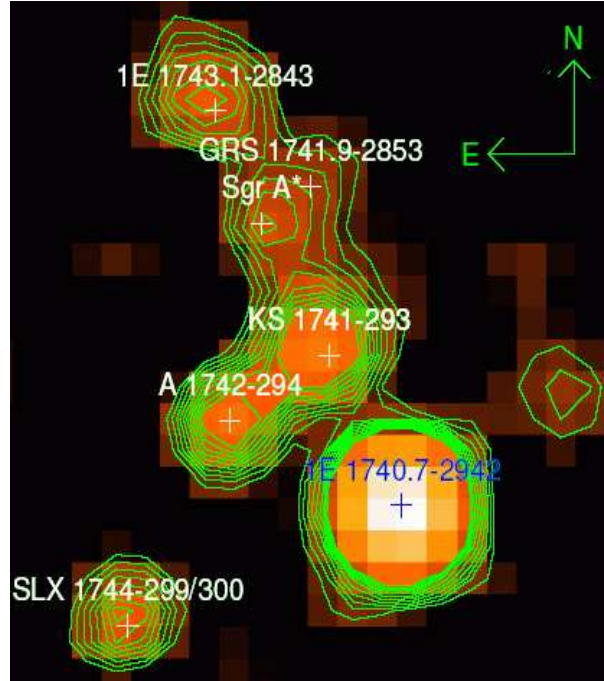


FIG. 1.6 – Observation d'un micro quasar par INTEGRAL, Belanger et al. (2003).

### Catalogues issus de relevés dans le domaine optique

Les relevés optiques sont les plus fréquents et les plus utilisés pour construire des catalogues de quasars. Ceci est essentiellement dû au fait que l'on récolte

plus d'informations dans ce domaine d'énergie et que les observations sont plus aisées car les technologies dans ce domaine sont mieux maîtrisées.

La recherche des quasars dans le domaine optique a ceci de pratique que les astronomes sont habitués à utiliser les filtres optiques. Néanmoins, dans le domaine optique les quasars, et NAG sont loin d'être les objets majoritaires ; tout au contraire, la plus grande partie des objets sont des galaxies et des étoiles, selon la région du ciel où l'on observe. Il s'agit donc de pouvoir distinguer le plus rapidement et le plus efficacement possible les quasars de tous ces objets. Une partie assez simple, si l'on s'intéresse uniquement aux quasars, est de supprimer du relevé, par critères morphologiques, les galaxies résolues. Néanmoins cela ne permet pas d'exclure étoiles et galaxies d'aspect ponctuel.

L'une des premières méthodes mises en oeuvre pour les grands relevés fut de distinguer les quasars par leur couleur particulièrement bleue. En effet, les couleurs de ces objets dans les filtres larges (couvrant plusieurs centaines d'angstroems), qui ne nécessitent donc que peu de temps d'observations, se distinguent assez facilement de celles des galaxies ou des étoiles. Je reviendrai plus en détail sur la façon de les distinguer par leur couleur dans la partie 3.3. D'une façon simple et pour introduire la motivation qui a conduit à la réalisation de ces relevés, on peut considérer une sélection d'objets tout à fait basique qui est réalisée sur un simple critère selon la couleur U-B. En effet, les étoiles de la séquence principale ont une couleur U-B de l'ordre de 0.5 magnitude alors qu'un noyau actif de galaxie typique, à bas décalage spectral, a une couleur U-B de l'ordre de -1 magnitude. Il est donc possible d'exclure une grande partie des étoiles, il existe bien sûr des contaminants telles les naines blanches. L'un des premiers relevés fut réalisé par Markarian en 1967. Un relevé d'importance fut le Palomar Bright Quasar Survey par Schmidt et Green (1983) ; leur sélection atteint une efficacité d'approximativement 7% , avec 114 NAG confirmés spectroscopiquement. Il est clair que l'utilisation de cette seule couleur ne permet pas de rendre compte de tous les quasars, en tout cas de toutes leurs caractéristiques, et donc par une telle sélection les relevés manquent un nombre important de quasars qui ont probablement des couleurs différentes. Je reviendrai plus tard sur les effets de sélections, liés à l'efficacité. D'autres relevés ont essayé très tôt d'utiliser comme critère de sélection une autre caractéristique importante des quasars : leurs larges raies d'émission dans l'optique et l'ultraviolet.

La sélection par critères de couleur nécessite de toute façon une confirmation spectroscopique, au final il y a donc une perte de temps d'observation. Une solution qui fut adoptée fut de placer un système de dispersion devant les détecteurs, ce qui permet d'obtenir une image contenant une myriade de spectres à très basse résolution. Ainsi en plus de critères sur les couleurs, on peut combiner la détection de raies d'émission grâce à la spectroscopie sans fente. Cette technique s'applique essentiellement aux quasars à un décalage spectral supérieur à 2, là où la raie d'émission Lyman- $\alpha$  (la plus importante) rentre dans le domaine visible (voir parmi les premiers relevés Hoag et Smith 1977). L'un des problèmes qui émergea fut l'identification des raies d'émission directement sur les détecteurs. En effet, la recherche dans l'image d'un pic d'émission se faisait à l'oeil et varie d'un observateur à l'autre. Avec les relevés plus récents, une sélection automatique fut mise en place. Dans le Large Bright Quasar Sample par

exemple (Foltz et al. 1987), qui permet l'identification d'à peu près 1000 quasars.

L'un des très récents sondages qui permet une avancée importante dans la base de connaissance des quasars est le 2dF QSO redshift Survey. Celui-ci a permis de sélectionner par critères de couleur ( $u$ ,  $b_j$ ,  $r$ ) 44 576 objets dont 23 338 se révèlent être des quasars. Ce relevé a été complété par le 6df QSO redshift survey qui identifie 322 quasars en plus. Les données actuelles proviennent désormais du SDSS (Sloan Digital Sky Survey) qui a pour objectif de découvrir plus de 100 000 quasars avec une méthode de sélection par les couleurs; je décrirai plus longuement celui-ci dans la partie 2.1.

La sélection à partir de l'ultra-violet n'a pas été l'objet de grands relevés, pas suffisamment sensible en tout cas. Tout récemment a été lancé le premier satellite qui va permettre de créer ce premier catalogue : GALEX (Galaxy Evolution Explorer lancé par la NASA en 2003). La mission a pour but de faire le relevé de plus de 2000 degrés carrés dans l'ultra-violet (1350-1800 Å et 1800-2800 Å), jusqu'à une magnitude de 25.5. Le but premier de GALEX n'est pas de rechercher les NAG, mais il se trouve que cette mission est parfaitement adaptée pour en étudier les propriétés dans l'ultra-violet, essentiellement en faisant une corrélation avec les catalogues de NAG connus.

Au final la description complète des caractéristiques des quasars passe par l'utilisation de grands échantillons. Les premières recherches de quasars ne pouvaient que difficilement utiliser le peu de connaissance que l'on avait sur ces objets. Néanmoins les derniers efforts dans ce domaine ont permis de tirer profit des données acquises; je reviendrai en détail sur ce point avec la description du SDSS. Mais il ne faut pas être ségrégationniste par rapport aux domaines d'énergie et ne rien oublier pour être sûr d'avoir une vue panchromatique des quasars et de leur propriétés.

### 1.2.2 Méthodes de recherche suivant les différents domaines d'énergie

Quelles sont les différentes méthodes qui ont été utilisées pour rechercher les quasars jusqu'à aujourd'hui? Dans un premier temps on a vu que les quasars ont été recherchés via des observations radio, mais l'efficacité est faible, dû à la faible quantité de quasars émettant effectivement en radio. Ensuite, d'après leur spectre on a pu comprendre que ces objets présentaient un excès d'émission dans le bleu comparé à la plupart des étoiles. Il s'agissait alors de faire la recherche via des excès simples de couleur, essentiellement dans le bleu. Des excès U-R; U-G, qui ne font jamais que sonder une infime partie des caractéristiques des quasars, ont permis de débiter des recherches à grande échelle. La spectroscopie sans fente a permis d'enrichir les catalogues de quasars du fait de la facilité de repérer les raies d'émission même sur des spectres basse résolution. La recherche utilisant les données optiques peut se baser aussi sur la variabilité des quasars. Vinrent ensuite les recherches des quasars dans d'autres domaines, X, infra-rouge, rayon gamma; plus pour comprendre la nature des quasars que pour agrandir les catalogues. L'astrométrie a aussi été utilisée pour la recherche de quasars, du fait de leur absence de mouvement propre.

## Recherche dans le domaine optique

Les grands sondages du ciel utilisent plusieurs filtres couvrant une grande partie du domaine optique, ce qui permet de travailler dans ce qu'il convient d'appeler l'espace des couleurs du sondage, c'est-à-dire l'espace constitué par les différences entre les magnitudes calculées dans les filtres utilisés par le sondage. Dans cet espace les objets peuvent se différencier par un excès dans une couleur ou dans l'autre. Les méthodes de recherche dans le domaine optique peuvent se départager en deux grands groupes. Le premier correspond à l'utilisation des plans dans l'espace des couleurs qui permet de définir des régions où il faut sélectionner des objets ou bien les exclure. Par exemple pour sélectionner les quasars dans les diagrammes couleur-couleur du 2dF. La sélection en couleur  $u$   $bj$   $r$  est faite en prenant les objets répondant aux critères de couleur (représentés figure 1.7) :

$$u - bj \leq -0.36 \quad (1.1)$$

$$u - bj < 0.12 - 0.8(bj - r) \quad (1.2)$$

$$bj - r < 0.05 \quad (1.3)$$

$$(1.4)$$

Le second groupe consiste à chercher la meilleure adaptation possible d'un spectre du type de l'objet recherché aux couleurs des objets du catalogue. En clair, il s'agit de comparer les magnitudes théoriques obtenues à partir des informations spectrales de différents objets (étoiles, quasars, galaxies) aux magnitudes observées. Ceci implique que l'on suppose connaître tous les spectres possible dans l'univers et il faut ensuite calculer la probabilité pour chacun d'eux de reproduire les couleurs observées.

Cette méthode a notamment été appliquée dans divers cas :

- Le relevé COMBO 17 (Wolf et al. 2004). "Adapter des distributions d'énergie spectrale revient à adapter un spectre haute résolution à un spectre faible résolution". La partie faible résolution étant en fait le catalogue photométrique obtenue à travers divers filtres. Par exemple COMBO-17 utilise 17 filtres pour sa photométrie de 3500 à 9300 Å ce qui correspond grossièrement à un spectre de très faible résolution. Le champ couvert est de 31.5' x 30' à une profondeur en R d'à peu près 24. Il est bon néanmoins de signaler que l'utilisation d'un si grand nombre de filtres permet d'atteindre une plus grande précision dans l'adaptation aux modèles, mais consomme beaucoup plus de temps de télescope.
- Cette alternative a aussi été choisie par le relevé BATC (Beijing-Arizona-TAipei-Connecticut) pour rechercher des quasars et déterminer leur décalage spectral (Wu et al. 2003).
- Une autre méthode de recherche a été mise au point par l'équipe de GAIA, qui avait prévu d'utiliser des filtres pour faire le même travail de classification. L'équipe avait alors comparé trois types de méthodes pour sélectionner les objets (cf. Claeskens et al. 2006) : une méthode d'adaptation à des distributions d'énergie spectrale, un arbre de décision et un réseau de neurones. Leur conclusion, semble favoriser la deuxième méthode.



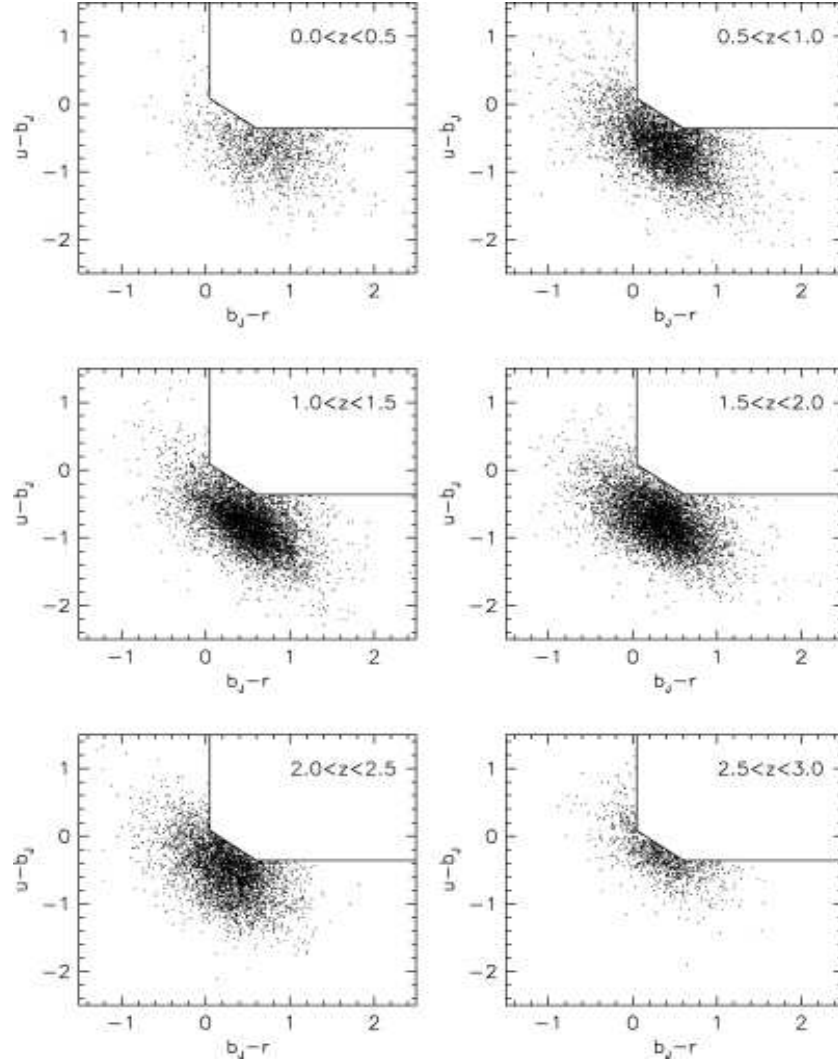


FIG. 1.7 – Diagramme couleur-couleur pour les quasars du 2QZ/6QZ arrangé en échantillon de décalage spectral de  $\Delta z = 0.5$ . Les lignes indiquent les limites de la sélection, les points sont les quasars présents dans chaque intervalle de décalage spectral indiqué en haut à droite de chaque panneau. A bas décalage spectral ( $0.0 < z < 0.5$ ) les couleurs des quasars sont plus rouges en  $b_J - r$ , donc la sélection en  $u - b_J$  a plus d'impact. A grand décalage spectral ( $z > 2$ ) les couleurs deviennent de plus en plus rouges. Croom et al. (2004 papier XII).

A noter aussi, bien qu'il ne s'agisse pas d'un relevé directement dédié à la recherche de quasars le VVDS (VIMOS VLT Deep Survey O. Le Fèvre et al. 2004) qui est un relevé spectroscopique. Celui-ci couvre des champs de Galex, l'étude de la nature des objets pourra donc s'enrichir de ce relevé en ultra-violet.

### Recherches venues d'ailleurs

Les autres domaines d'énergie sont essentiels pour la compréhension pan-chromatique des quasars. La sélection via le domaine radio s'est généralement effectuée en parallèle, afin d'être plus efficace, avec un relevé optique en utilisant des excès de couleurs simples comme U-B. En infra-rouge, il est possible d'essayer de rechercher l'excès de luminosité dans ce domaine d'énergie (Low et al. 1988).

Les quasars peuvent encore être distingués des étoiles par leur absence de mouvement propre (Sandage & Luyten 1967, Kron and Chiu 1981, Koo et al. 1986). Dans ce dernier cas la sélection n'est pas très efficace et doit être utilisée en combinaison avec d'autres telles que la variabilité.

Techniquement il est plus facile de prévoir une méthode de sélection principale pour un relevé, et ensuite d'utiliser "en bonus" ce que peuvent nous apporter les autres méthodes. Typiquement, le SDSS a sélectionné ses candidats sur des critères de couleur, tout en prenant en compte les sources radio de FIRST. Ceci permet de limiter les effets de sélection d'une part et en plus d'améliorer l'efficacité du relevé. Ceci ne coûte pas plus cher en temps de télescope car les deux se complètent et peuvent être dédiés à des études différentes ; l'important pour pouvoir les exploiter en parallèle étant qu'il soit possible de faire une corrélation entre les deux catalogues, et donc que les mêmes régions du ciel soient couvertes. Il s'agit alors de faire la corrélation entre catalogues, c'est-à-dire d'identifier chaque détection dans chaque catalogue à une source dans le ciel. L'une des méthodes les plus prometteuses étant sûrement la corrélation entre catalogues optiques et X. La précision en X est désormais suffisante pour que la corrélation de position soit réalisable sans de trop fortes ambiguïtés.



## Chapitre 2

# Les sondages utilisés : CFHTLS et SDSS

*"We're so busy watching out for what's just ahead of us that we don't take time to enjoy where we are."*

– Calvin

## Sommaire

<b>2.1</b>	<b>SDSS</b>	<b>38</b>
2.1.1	La photométrie du SDSS	39
2.1.2	Échantillons du SDSS	42
	Base Spectroscopique	44
	Base Photométrique	45
	Catalogue de quasars	47
<b>2.2</b>	<b>CFHTLS</b>	<b>47</b>
2.2.1	Généralités - Traitement et accès aux données	47
2.2.2	Photométrie - Comparaison avec le SDSS	49
<b>2.3</b>	<b>Autres champs observés au CFHT</b>	<b>56</b>

Mon travail s'est porté sur la partie optique de l'émission des quasars. J'ai entre autres utilisé les données de deux sondages pour y effectuer ma recherche. D'un coté le SDSS, Sloan Digital Sky Survey, qui représente une avancée énorme en ce qui concerne la quantité de données (galaxies et quasars). De l'autre le CFHTLS (Canada France Hawaï Telescope Legacy Survey) qui représente un champ quasi vierge en ce qui concerne la recherche de quasars.

De nouvelles méthodes de recherche de quasars ont été développées pour le SDSS. Ce relevé fournit un grand nombre de spectres de quasars et représente une avancée statistique importante par rapport aux précédents relevés. Le SDSS sera donc utilisé pour tester ma méthode de sélection qui sera mise en place pour les champs du CFHTLS. C'est-à-dire que je vais utiliser un champ suffisamment grand du SDSS contenant des objets observés en spectroscopie. J'appliquerai alors sur les catalogues photométriques du SDSS ma sélection, et je pourrai déterminer si celle-ci retrouve les quasars présents dans le SDSS. La migration de ma sélection vers le relevé CFHTLS sera possible grâce aux ressemblances entre les jeux de filtres utilisés dans les deux sondages, qui font que les méthodes de sélection optimale sont très proches pour les deux relevés. Je vais décrire en détail ces deux relevés, notamment les différences en photométrie.

## 2.1 SDSS

Le SDSS a pour but de couvrir une très grande partie du ciel pour permettre la cartographie de près de 100 millions d'objets en photométrie et de relever les spectres de près de 100 000 quasars et d'un million de galaxies. Les observations de ce relevé ont la particularité d'être effectuées avec une caméra CCD et donc nous avons la possibilité d'en disposer directement sous forme numérique. Le sondage opérera durant cinq années et couvrira dans le ciel 45 bandes continues séparées de 2,5 degrés dans l'hémisphère nord et trois bandes non-continues à -10, 15 et 0 degré dans l'hémisphère sud. Le relevé couvrira ainsi près de 10 000 degrés carrés dans l'hémisphère nord et 740 dans le sud (cf. figure 2.1).

La caméra utilisée est constituée de trente CCD (Charge-Coupled Device) organisés en six colonnes de cinq CCD chacune qui sont couverts par un filtre différent :  $u$ ,  $g$ ,  $r$ ,  $i$  et  $z$ . (Fukugita et al. 1996).

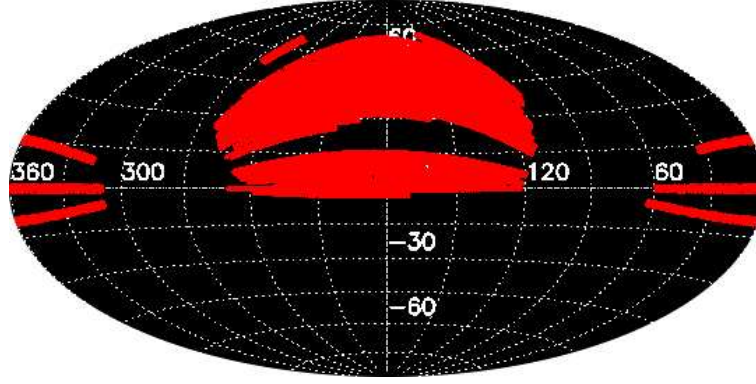


FIG. 2.1 – Représentation de la couverture du relevé photométrique du SDSS au “Data release 4” en coordonnées équatoriales. Issue du site <http://www.sdss.org>

Le SDSS utilise deux spectrographes multi-objets chacun comprenant 320 fibres optiques ; deux caméras permettent de récupérer les images, l’une dans le bleu de 3900 Å à 6000 Å et l’autre dans le rouge de 5800 Å à 9100 Å. Les possibilités d’études grâce à ces données sont multiples, que ce soit les objets proches de la Terre, ou les naines brunes dont celles de classe T (la seconde naine de ce type recensée a été découverte dans le SDSS par Fan et Strauss 2000), que ce soit pour l’observation du halo galactique pour y découvrir les régions susceptibles d’abriter des étoiles et découvrir comment elles ont été arrachées à notre galaxie, ou pour l’étude des structures à grande échelle à partir de la détection de galaxies, ou encore la recherche des quasars, dont la découverte est l’un des principaux but du sondage, notamment ceux à grand décalage spectral (les quasars situés à  $z > 5$  ayant été pour la plupart découverts dans le SDSS). L’ensemble des données relatives à ce relevé sont disponibles sur les sites internet : <http://www.sdss.org/> et <http://www.sdss.org/dr4/index.html>.

### 2.1.1 La photométrie du SDSS

Le SDSS a mis en place un système photométrique particulier, basé sur des magnitudes calculées en utilisant la fonction sinus hyperbolique au lieu d’une fonction logarithmique (Lupton et al. 1999). En effet, les magnitudes habituelles utilisent une échelle logarithmique pour mesurer les flux, ce qui est particulièrement utile pour pouvoir observer de grands intervalles d’énergie, et les couleurs relatives s’obtiennent très facilement par des différences de magnitudes. Par exemple les magnitudes de type  $AB$ , définies par Oke & Gunn (1983) :

$$AB = -2.5 \log(f_\nu) + 48.6 \quad (2.1)$$

avec  $f_\nu$  le flux en  $\text{ergs cm}^{-2} \text{s}^{-1} \text{Hz}^{-1}$ . Ces avantages sont effectifs surtout pour des objets très lumineux. Mais lorsque le flux mesuré est proche du fond de ciel et que le bruit instrumental est important, les magnitudes mesurées sur une telle échelle ont des comportements singuliers, dûs à l’approche du flux nul. Le cas qui se rencontre le plus régulièrement est la détection d’un objet dans un des filtres mais la mesure du flux dans un autre filtre est quasiment impossible. Il

faut donc traiter toutes les non-détections différemment, en leur attribuant par exemple des valeurs limites. L'une des solutions consiste à utiliser une échelle d'énergie linéaire, par exemple les Janskys, mais alors les études basées sur les rapports de flux deviennent peu pratiques (car les rapports de flux ne correspondent plus à des différences de magnitude). L'équipe du SDSS, en particulier Lupton et al. (1999) proposent une alternative à la définition des magnitudes en préservant les avantages connus et en évitant (sic) leurs inconvénients. Ils remplacent la fonction logarithmique par la fonction sinus hyperbolique, celle-ci devient logarithmique pour de grandes valeurs et est linéaire près de l'origine.

$$\sinh^{-1}(x) = \ln \left[ x + \sqrt{x^2 + 1} \right] \rightarrow \begin{cases} \operatorname{sgn}(x) \ln |2x|, & \text{if } |x| \gg 1 \\ x, & \text{if } |x| \lesssim 1 \end{cases} \quad (2.2)$$

où  $x \equiv f/f_0$  est le flux normalisé qui peut décrire la magnitude apparente usuelle  $m$  comme :

$$m \equiv -2.5 \log_{10} x = -(2.5 \log_{10} e) \ln x \equiv -a \ln x. \quad (2.3)$$

avec  $f_0$  le flux d'un objet avec une magnitude de 0.0, et  $a \equiv 2.5 \log_{10} e = 1.08574$  est le rapport de Pogson (Pogson 1856). En définissant la magnitude (asinh magnitude)  $\mu$  comme étant :

$$\mu(x) \equiv -a \left[ \sinh^{-1} \left( \frac{x}{2b} \right) + \ln b \right] \quad (2.4)$$

avec  $a$  et  $b$  des constantes.  $b$  est une constante permettant de déterminer le flux à partir duquel le comportement linéaire débute. Noter le comportement asymptotique de  $\mu$  pour des valeurs grandes et petites de  $x$  :

$$\lim_{x \rightarrow \infty} \mu(x) = -a \ln x = m \qquad \lim_{x \rightarrow 0} \mu(x) = -a \left[ \frac{x}{2b} + \ln b \right]. \quad (2.5)$$

Ainsi pour  $x \rightarrow \infty$ ,  $\mu$  s'approche de  $m$ , et ce pour n'importe quel choix de la valeur de  $b$ . D'un autre côté lorsque  $|x| \lesssim b$ ,  $\mu$  a un comportement linéaire en  $x$  ; et ainsi on revient graduellement au comportement logarithmique (cf figure 2.2). La valeur de  $b$  est reliée au bruit typique du ciel à 1 sigma dans une PSF (point spread function) avec un seeing de 1", et ce relativement à  $f_0$ . La magnitude limite est donc  $-a \ln b$ .

Pour chaque champ il faut déterminer : le point zéro, le coefficient d'extinction et la masse d'air.

Le point zéro est défini comme étant la magnitude d'une étoile qui produit un taux de coup (un coup est aussi appelé 'Analog to Digital Unit', ADU) par seconde égal à 1 dans un filtre donné. Le point zéro permet de définir les magnitudes utilisées ; pour le SDSS il est déterminé à partir d'étoiles standards et est précis à mieux que de 2%. La calibration en magnitude nécessite tout d'abord de déterminer le rapport entre le taux de coups observé corrigé de l'extinction et  $f_{pz}$ , le taux de coup pour le point zéro.

Il faut pour cela :

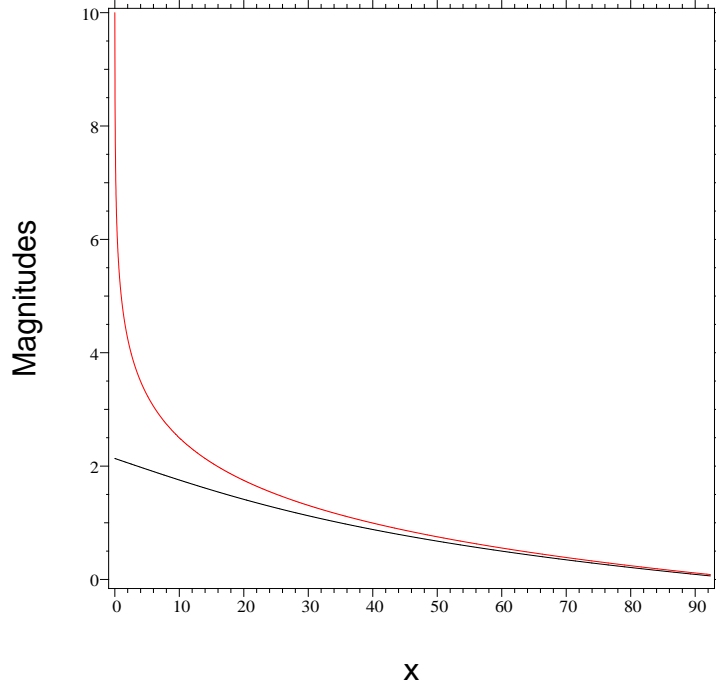


FIG. 2.2 – Comparaison des magnitudes logarithmiques en rouge et des magnitudes asinh en noir, pour des valeurs arbitraires de  $b$  et de  $x$ .

- Convertir le nombre de coups observés en un taux de coups par seconde, à partir du temps d'exposition.
- Corriger ce taux de l'extinction atmosphérique en utilisant le coefficient d'extinction  $k$  et la masse d'air (*airmass*)
- Diviser par le taux de coup pour le point zéro qui est donné par  $f_{pz} = 10^{-0.4*z_p}$  coups par seconde avec  $z_p$  étant le point zéro.

Soit au final :

$$f/f_{pz} = \text{coups}/t * 10^{0.4*(z_p + k*airmass)} \quad (2.6)$$

avec  $t$  le temps d'exposition. Et ensuite calculer la magnitude pour  $f/f_{pz}$  :

$$m_s = -\frac{2.5}{\ln(10)} * [\text{asinh}(\frac{f}{f_{pz}}) + \ln(b)] \quad (2.7)$$

et l'erreur associée :

$$\text{erreur}(m_s) = \frac{2.5}{\ln(10)} * \frac{\text{erreur}(\text{coups})}{t} * \frac{1}{2b} * \frac{10^{0.4*(z_p + k*airmass)}}{\sqrt{1 + [\frac{f}{2b}]^2}} \quad (2.8)$$

L'erreur sur les coups est déterminée comme suit :

Le nombre de photo-électrons reçus qui sont requis pour produire un coup est déterminé par le gain du détecteur :

$$\text{photoelectrons} = \text{coups} * \text{gain} \quad (2.9)$$



filtre	b	Zéro-Flux Magnitude [ $m(f/f_0 = 0)$ ]	$m(f/f_0 = 10b)$
u	$1.4 \times 10^{-10}$	24.63	22.12
g	$0.9 \times 10^{-10}$	25.11	22.60
r	$1.2 \times 10^{-10}$	24.80	22.29
i	$1.8 \times 10^{-10}$	24.36	21.85
z	$7.4 \times 10^{-10}$	22.83	20.32

TAB. 2.1 – SDSS - Les valeurs du paramètre  $b$  sont fixés approximativement à 1 sigma du bruit du ciel. La troisième colonne donne la magnitude  $-\ln(b)$ . La dernière colonne donne la magnitude asinh associée à un objet pour lequel  $f/f_0 = 10b$

L'estimation des erreurs doit tenir compte du bruit poissonien de la contribution de l'objet plus celle du ciel sur la même surface ; doit aussi rentrer en compte le bruit de lecture qui est estimé pour chaque observation et est noté comme 'dark variance' auquel il faut ajouter l'erreur sur l'estimation du ciel dans l'image (relatif à la contribution du ciel au flux total) :

$$erreur(coups) = \sqrt{[coups + ciel]/gain + Npix * (darkvariance + cielErr)} \quad (2.10)$$

avec 'coup' le nombre de coups pour l'objet, 'ciel' est le nombre de coups pour le ciel comptés sur une surface identique à celle de l'objet. Npix est l'aire couverte par l'objet en pixels, gain et 'dark variance' sont comme définis précédemment et cielErr est l'erreur estimée sur la moyenne du niveau du ciel dans l'image. Les paramètres  $b$  et le point zéro sont donnés dans la table 2.1.

La base de données du SDSS donne accès à différents types de magnitude selon la façon dont le flux est extrait de l'image. Il est important pour la suite de noter que pour des objets ponctuels la meilleure façon de déterminer le flux total est d'adapter un modèle de PSF 'point spread fonction'. Cette magnitude a été déterminée en décalant l'image d'un objet de façon à le centrer parfaitement sur un pixel et ensuite d'y adapter un modèle gaussien pour la PSF. Certaines corrections sont appliquées à partir d'étoiles brillantes qui permettent de réduire la dépendance des magnitudes en fonction du 'seeing'.

De plus lors de l'application du dérougissement galactique il est nécessaire de tenir compte des erreurs sur l'estimation de celui-ci. Typiquement il faut rajouter 15% de la valeur de dérougissement en quadrature aux erreurs photométriques.

### 2.1.2 Échantillons du SDSS

Le SDSS va m'être utile pour étalonner et tester mes outils de sélection. Dans un premier temps il est nécessaire de définir quelles sont les parties de ce relevé que je vais utiliser, quelles sont les informations à extraire qui vont m'être utiles pour conduire et évaluer les tests.

Comme dit précédemment la base de données se découpe en deux grandes parties, une spectroscopique et l'autre concernant l'imagerie que j'appellerai photométrique (cf. figure 2.3).

Il m'a fallu définir les échantillons du SDSS sur lesquels il fallait conduire les tests de sélection. L'idée dans ce cas est de pouvoir disposer d'échantillons

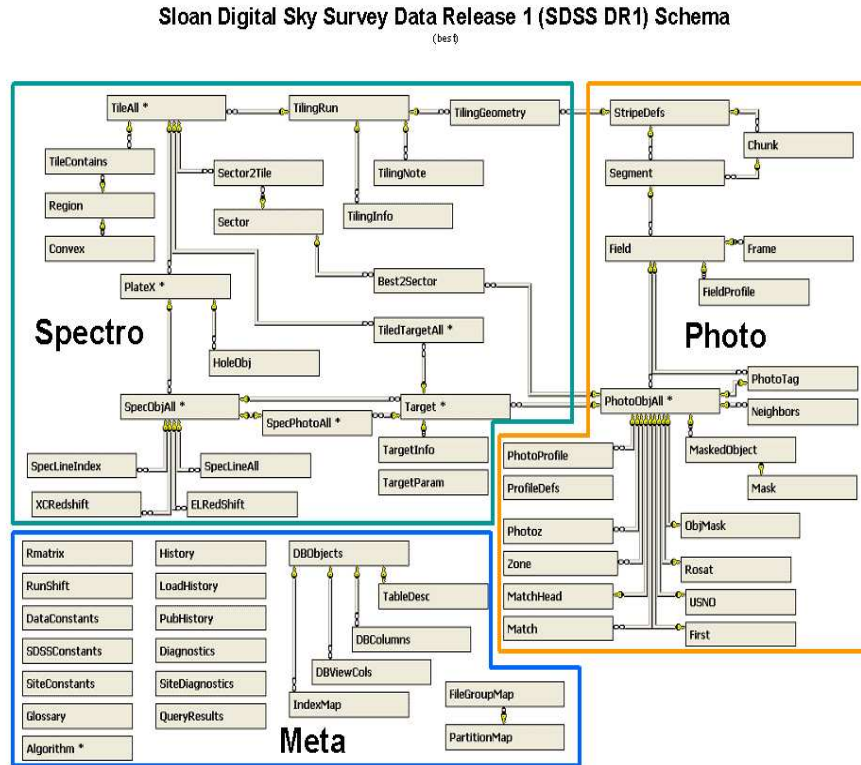


FIG. 2.3 – Schéma de la base de données pour le SDSS telle qu'elle a été conçue pour le "data release I". On remarque qu'il y a principalement deux bases de données interconnectées : la photométrie et celle spectroscopique. Celles-ci contiennent des informations complémentaires. Issue du site <http://sdss.org>

d'objets les moins biaisés possible (c'est-à-dire ayant un effet de sélection minimum). Certainement, l'absence de biais n'est pas possible puisque le SDSS fait l'objet de sélection et de critères particuliers. Ces biais sont notamment liés à l'absence ou à la sous-représentation de quasars obscurcis, puisque la sélection se fait essentiellement dans le domaine optique. D'après la description de la base de données du SDSS, il est possible de travailler à partir de deux bases particulières.

La base spectroscopique, qui contient un maximum d'information puisque les objets qui y sont présents ont fait l'objet d'un suivi spectroscopique. Néanmoins cette base est aussi celle qui est soumise aux plus forts biais de sélection. En effet, les objets présents dans cette base ont fait l'objet d'une sélection par l'équipe du SDSS. L'utilisation d'échantillons issus de cette base ne va donc pas nous permettre d'estimer les contaminants dû à notre sélection.

La base photométrique contient l'ensemble des objets tombant sous la limite de détection du relevé. Cette base a l'avantage de n'avoir aucun biais de sélection, mais l'information disponible sur les objets y est limitée. En effet, ils ne disposent pas tous d'une identification spectrale, il sera donc impossible de dire

lorsque l'un de nos candidats fera partie des objets sans identification spectrale si c'est un 'bon' candidat ou pas. Je reviendrai sur ce point et sur le traitement appliqué dans ces cas (cf. section 4.2.2).

A ces échantillons s'ajoute la liste exhaustive des quasars dans le SDSS : le catalogue de quasars publié par Schneider et al. (2003, 2005). En effet, la classification dans la base de donnée en ligne se fait de façon tout à fait automatique et est bien sûr sujette à des erreurs de classification ou à des omissions. C'est pourquoi une recherche impliquant une inspection manuelle et souvent une révision de la classification automatique a été effectuée par Schneider et al. (2003 et 2005). Sur le DR1 (data release 1) en 2003 et sur le DR3 (data release 3) en 2005. Cet échantillon sera utile pour travailler sur un grand nombre de quasars, incluant des types de quasars peu représentés ou difficiles à reconnaître de façon automatique, puisque même à partir de leur spectre, la classification automatique du SDSS s'est trompée sur certains d'entre eux.

Je vais décrire sur quels critères j'ai sélectionné les échantillons d'objets dans chacune des bases ainsi que les données que j'ai extraites du site du SDSS.

## Base Spectroscopique

Pour un champ d'observation typique le SDSS dispose de 640 fibres (2 spectrographes de 320 fibres) sur un peu plus de 6 degrés carré ; une trentaine sont réservés pour le ciel, près de vingt pour des étoiles ou des naines standards ; près de 420 fibres sont dédiées aux galaxies et approximativement 60 aux galaxies lumineuses rouges (GLR) et en gros 80 sont réservées au quasars. Ce qui représente une moyenne de 18 observations de candidats quasar par degré carré et la sélection de Richards et al. (2002) (détaillée chapitre 3) donne une moyenne de 15 candidats par degré carré.

Parmi les quelques fibres restantes, sont observés par ordre d'importance : des étoiles (une quinzaine de fibre) sur des critères de couleurs (essentiellement des étoiles bleues du halo, des naines brunes, blanches et rouges et des variables cataclysmiques), des cibles 'au hasard' (appelées 'serendipity target' entre 7 et 8 fibres) qui ne répondent pas aux critères de sélection automatiques et enfin des cibles issues du relevé ROSAT (5 fibres). Cette sélection cible donc essentiellement les quasars (parmi les objets ponctuels), ou les objets éloignés du locus des étoiles.

J'ai défini mes échantillons sur la base spectroscopique de la façon suivante :

Les données qui me seront nécessaires sont l'identifiant du quasar (SpecObjID), son décalage spectral reconnu à partir du spectre de façon automatique afin de le comparer plus tard au décalage déterminé grâce à sa photométrie. Il me faut aussi récupérer ses magnitudes dans chaque filtre à travers la fibre, et sa position afin de pouvoir l'identifier. Ensuite, la classification automatique mise en place par le SDSS discerne les quasars avec un drapeau 'specClass' égal à 3 et les quasars à grand décalage spectral avec un drapeau 'specClass' égal à 4. Enfin, le SDSS a mis en place, à partir des erreurs sur la mesure du décalage spectral et d'éventuels problèmes avec les raies identifiées un drapeau ('NOT\_QSO') signalant que l'objet n'est finalement pas un quasar. Le tout est fourni par la requête SQL (Structured Query Language) présentée ci-dessous :

```

select SpecObjID, z, fiberMag_u, fiberMag_g, fiberMag_r, fiberMag_i, fiberMag_z, ra, dec
from SpecObj p
where p.ra>214.35
AND
p.ra<215.3
AND
p.dec>52.17
AND
p.dec<53.3
AND (specClass = 3 OR specClass = 4)
AND (zWarning & dbo.fSpecZWarning('NOT_QSO')) = 0

```

La première ligne représente les données à extraire de la table SpecObj. La dernière requête SQL est apparue essentielle afin d'éviter d'avoir des étoiles dans l'échantillon de quasars. Néanmoins ce garde fou n'est pas complètement sûr, car il y a des objets présentant le drapeau ('NOT\_QSO') qui sont en réalité des quasars. Par exemple l'objet ayant pour identifiant 'SDSS J112253.49+005329.8' qui est tout d'abord classé comme quasar (i.e. ayant 'specClass' = 3 ou 4) mais qui possède ce drapeau le signalant comme n'étant pas un quasar (drapeau 'NOT\_QSO') apparaît d'après l'observation de son spectre dans la base de données, être bel et bien un quasar à un décalage spectral de 4.55. Ce qui signifie que cette base de quasar n'est pas complète. Ceci est exactement le type de cas rencontré par Schneider et al. (2003, 2005) qui nécessite un classement visuel et 'à la main' pour être complet. Je n'ai pas réalisé ce classement visuel, et j'ai donc négligé les quasars ainsi manqués.

Pour déterminer quelles régions sont les plus complètes au niveau du suivi spectroscopique (c'est-à-dire dont les objets du champ devant être ciblés l'ont bien tous été), je me suis basé sur l'étude de la complétude de Vanden Berk et al. (2005) ou à partir de la partie 'navigate' du site du SDSS qui permet de repérer les régions suivies en spectroscopie.

## Base Photométrique

La sélection dans la base photométrique a été faite suivant plusieurs critères ; la plupart sont issus de la description qu'en fait Richards et al. (2002). En effet, certains objets dont la photométrie n'est pas digne de confiance doivent être exclus de nos échantillons, ceux qui sont trop brillants (drapeau 'BRIGHT' en SQL), ceux qui sont saturés ('SATURATED'), qui sont sur le bord des détecteurs ('EDGE'), et enfin ceux qui ont de trop proches voisins pour en être correctement séparés sans influencer leur photométrie ('BLENDED'). Les objets dis 'BRIGHT' correspondent à des objets qui ont été détectés à plus de 200 sigma dans la bande  $r$  sur un ciel estimé de façon global ; leur photométrie a ensuite été réévaluée avec un ciel local. La première estimation se voit affecté le drapeau 'BRIGHT'. Les objets sélectionnés doivent avoir été détectés à un niveau de confiance d'au moins  $5\sigma$  dans au moins une des bandes. Il peut y avoir des recouvrements entre chaque prise d'image et donc un objet peut être observé plusieurs fois. Afin d'éviter des objets qui apparaissent plusieurs fois dans l'échantillon il faut se restreindre à leur observation principale ('OK\_SCANLINE' et 'PRIMARY'). De plus je me suis, dans la plupart des cas, limité à une magnitude en  $i$  entre 15 et 20.2 qui est la limite de sélection de l'algorithme de reconnaissance mis en place par le SDSS, sauf dans certain cas particuliers sur lesquels je reviendrai. L'ensemble des contraintes de sélection permet de s'assurer que je traite des objets dont la photométrie est fiable.

Cette base va me permettre de construire les catalogues d'objets photométriques qui seront utilisés en entrée de l'outil de sélection HyperZ. Tout d'abord les données nécessaires sont les magnitudes dites "magnitudes psf" car mieux adaptées aux objets ponctuels (cf. 2.1.1). Il est également nécessaire de récupérer l'erreur photométrique qui y est associée, la position de l'objet pour son identification, le 'type' qui est un drapeau de classification selon leur morphologie (le drapeau 3 étant exclu car lié aux galaxies), l'identifiant de l'objet dans la base de données (afin de faciliter l'accès aux données pour d'éventuelles vérifications), les valeurs de l'extinction déterminée d'après la carte de Schlegel et al. (1998), et son éventuel identifiant dans la base spectroscopique (qui permet à la fois un accès à ses caractéristiques et une vérification de sa nature via son spectre). Il est important de noter que pour toutes les sélections dont je parlerai par la suite, tous les candidats issus de cette base de données qui n'ont pas d'identification spectrale seront considérés comme n'étant pas des quasars. Leur nature dans le SDSS est inconnue (car sans spectre), mais certaines caractéristiques permettent d'émettre quelques hypothèses, notamment des drapeaux 'TARGET\_QSO\_FAINT' qui les présentent comme étant des candidats quasars pour un suivi spectroscopique plus profond. La requête SQL donne :

```
select psfMag_u, psfMag_g, psfMag_r, psfMag_i, psfMag_z, psfMagErr_u, psfMagErr_g, psfMagErr_r, psfMagErr_i, psfMagErr_z, ra, dec, type, objID, extinction_u, extinction_g, extinction_r, extinction_i, extinction_z, SpecObjID
from PhotoObj p
where p.ra>214.35
AND
p.ra<215.3
AND
p.dec>52.17
AND
p.dec<53.3
AND
(flags_u & (dbo.fPhotoFlags('BRIGHT') + dbo.fPhotoFlags('SATURATED') +
dbo.fPhotoFlags('EDGE') + dbo.fPhotoFlags('BLENDED')))=0
AND
(flags_g & (dbo.fPhotoFlags('BRIGHT') + dbo.fPhotoFlags('SATURATED') +
dbo.fPhotoFlags('EDGE') + dbo.fPhotoFlags('BLENDED')))=0
AND
(flags_r & (dbo.fPhotoFlags('BRIGHT') + dbo.fPhotoFlags('SATURATED') +
dbo.fPhotoFlags('EDGE') + dbo.fPhotoFlags('BLENDED')))=0
AND
(flags_i & (dbo.fPhotoFlags('BRIGHT') + dbo.fPhotoFlags('SATURATED') +
dbo.fPhotoFlags('EDGE') + dbo.fPhotoFlags('BLENDED')))=0
AND
(flags_z & (dbo.fPhotoFlags('BRIGHT') + dbo.fPhotoFlags('SATURATED') +
dbo.fPhotoFlags('EDGE') + dbo.fPhotoFlags('BLENDED')))=0
AND(status & dbo.fPhotoStatus('OK_SCANLINE') > 0)
AND
((psfMagErr_u<0.2 AND psfMagErr_u>0.) OR (psfMagErr_g<0.2 AND psfMagErr_g>0.) OR
(psfMagErr_r<0.2 AND psfMagErr_r>0.) OR (psfMagErr_i<0.2 AND psfMagErr_i>0.) OR
(psfMagErr_z<0.2 AND psfMagErr_z>0.))
AND psfMag_i > 15.
AND psfMag_i < 20.2
AND type != 3
AND (status & dbo.fPhotoStatus('PRIMARY') > 0)
AND (status & dbo.fPhotoStatus('OK_STRIPE') > 0)
AND (status & dbo.fPhotoStatus('OK_RUN') > 0)
AND (status & dbo.fPhotoStatus('GOOD') > 0)
```

Pour déterminer les objets de cette base qui appartiennent aussi à la base spectroscopique il suffit de ne prendre que les objets qui ont fait l'objet d'un suivi c'est-à-dire avec un 'SpecObjID' différent de zéro en rajoutant la condition dans la requête SQL :

```
AND SpecObjID!=0
```

## Catalogue de quasars

Enfin, comme précisé plus haut, j'ai utilisé l'échantillon un peu particulier défini par Schneider et al. (2003, 2005). Celui-ci a été construit à partir des DR1 et DR3. Dans les deux cas une recherche a été menée de façon à récupérer tous les quasars présents dans la base spectroscopique. Les spectres ont été revus par des astronomes de façon visuelle et leur décalage spectral a lui aussi été revu. Il faut noter que ce catalogue ne se cantonne pas aux objets ponctuels, en effet nombre d'objets ne sont pas des quasars selon notre définition mais des NAG dont la galaxie hôte est résolue, que l'on peut qualifier de galaxie de Seyfert donc. J'ai pris ce catalogue tel quel, en y récupérant directement les données que j'utiliserai en entrée d'HyperZ.

## 2.2 CFHTLS

### 2.2.1 Généralités - Traitement et accès aux données

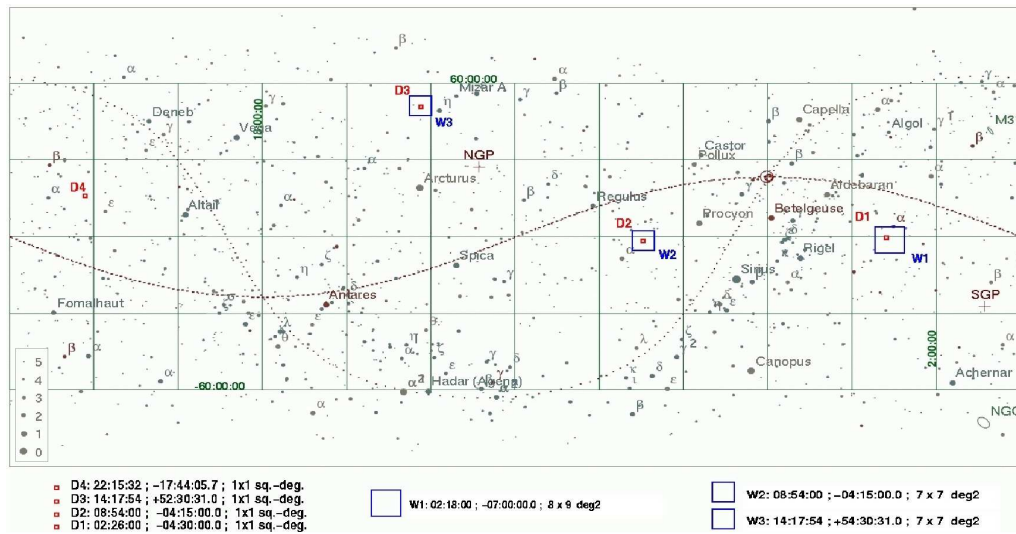


FIG. 2.4 – Position des champs 'Deep' et 'Wide' avec le plan de l'écliptique, ainsi que leur coordonnées. Issue du site Internet du CFHT

Le Canada France Hawaii Telescope Legacy Survey (CFHTLS) représente une avancée majeure dans le monde des relevés. Plus de 450 nuits réparties sur 5 années seront consacrées à ce relevé qui utilise une caméra CCD très large champ : MegaCam. Ce détecteur est composé de 36 CCD pour former un champ de vue de un degré carré. Le projet est divisé en trois relevés :

1. Le 'Very Wide', qui couvre une grande partie du plan de l'écliptique dans une bande de deux degrés carrés, pour un total de près de 1300 degrés carrés. Les images seront prises dans trois filtres ( $g$ ,  $r$  et  $i$ ) avec une magnitude limite en  $i$  de 24.0. Ce relevé permettra de créer un échantillon de données sur les objets de notre système solaire au-delà de Neptune.

Il permettra aussi de récolter des échantillons d'étoiles à des magnitudes limites bien plus faibles que celles du SDSS.

2. Le 'Wide', qui couvre 170 degrés carrés répartis en trois champs de 49 à 72 degrés carrés avec l'ensemble des filtres ( $u^*$ ,  $g'$ ,  $r'$ ,  $i'$ ,  $z'$ ), jusqu'à une profondeur de 24.5 en  $i$ . Ce relevé a pour but scientifique principal l'étude des grandes structures et de la distribution de matière à travers les effets de lentilles gravitationnelles faibles et la distribution de galaxies. Les observations en bande  $r$  se feront en deux phases séparées de trois ans, ce qui permettra également des relevés de mouvements propres.
3. Enfin le relevé 'Deep' qui couvre 4 degrés carrés dans quatre champs indépendants répartis dans le ciel de façon à avoir toujours la possibilité d'observer deux de ces champs durant à peu près n'importe quelle nuit de l'année. Les images seront réalisées avec tous les filtres disponibles ( $u^*$ ,  $g'$ ,  $r'$ ,  $i'$ ,  $z'$ ) avec comme limite en magnitude dans chaque filtre respectivement : 28.3, 28.5, 28.3, 28.3 et 27. Les échantillons de galaxies et de quasars récoltés vont permettre de poser de fortes contraintes sur l'évolution des objets. Grâce aux relevés à différentes époques on pourra détecter les mouvements propres d'objets ainsi que leurs variations en flux.

L'ensemble du CFHTLS est représenté figure 2.4. Les données acquises seront traitées suivant le diagramme 2.5. Les ellipses noires représentent les endroits où les utilisateurs pourront récupérer les données au CADC (The Canadian Astronomy Data Centre). Pour chaque étape de récupération est représenté le temps nécessaire après acquisition au télescope jusqu'à ce niveau de traitement. Les données que j'ai utilisées sont disponibles au niveau du processus de réduction par Elixir. Cette partie a pour but d'enlever les traces instrumentales des données. Ceci comprend l'utilisation de masques pour les pixels morts, correction des biais, des flat-fields, correction de franges dans les filtres  $i$  et  $z$ . Cette partie du traitement permet aussi d'extraire l'astrométrie à 0.2 arsec près. Un point très important est l'estimation du point zéro pour chaque filtre en utilisant les étoiles standards photométriques observées pendant l'acquisition de données. Le détail des processus de réduction peut être trouvé sur les pages du CFHT : <http://www.cfht.hawaii.edu/Instruments/Elixir/home.html>. Ces données réduites sont alors mises à la disposition des utilisateurs et sont prêtes pour passer à l'étape suivante, productrice des catalogues. Ces images sont ensuite fournies à l'équipe Térapix. Celle-ci a pour tâche de réaliser les images de poids, l'addition des données, le calcul d'une astrométrie précise, ainsi que la génération de catalogues de sources photométriques. Il est nécessaire d'attendre plusieurs mois pour pouvoir accumuler suffisamment de données et réaliser ces opérations. Ceci fait, les données réduites et calibrées sont délivrées à la communauté lors de 'release', nommés TXXXX, commençant à T0001 ; au moment où ces lignes sont écrites, le relevé fournit les données du T0003.

Le T0001 (données obtenues entre le premier juin 2003 et le 22 juillet 2004) fut accessible en novembre 2004 dont les caractéristiques sont disponibles sur

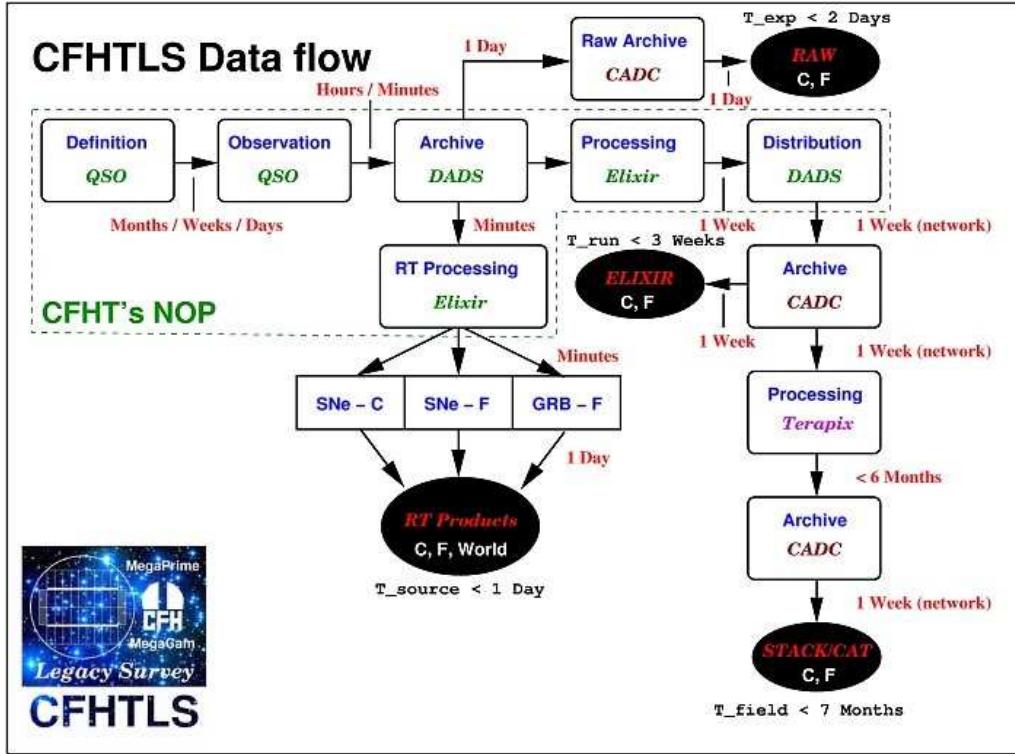


FIG. 2.5 – Cheminement des données du CFHTLS, et récupération par les utilisateurs (représentée par les ellipses). QSO étant ici l’acronyme pour Queued Service Observing. Issue du site du CFHT.

[http://terapix.iap.fr/article.php?id\\_article=382](http://terapix.iap.fr/article.php?id_article=382) et figure 2.6. Il est important de noter qu’à ce moment les données du champ D4 dans le filtre  $z$  étaient manquantes car les solutions astrométriques ne correspondaient pas à celles des autres filtres.

Le T0002 (données obtenues entre le premier juin 2003 et le 22 novembre 2004) fut révélé en juin 2005 dont les caractéristiques sont disponibles sur [http://terapix.iap.fr/article.php?id\\_article=198](http://terapix.iap.fr/article.php?id_article=198) et figure 2.7. Celui-ci ajoute suffisamment de données pour utiliser tous les champs, notamment le D4. De plus les magnitudes point zéro sont données relativement à celles du champ Deep 2 lorsqu’au moins 3 filtres sont disponibles sinon la référence est fixée à 30. Notez toutefois sur la figure 2.7 le temps d’exposition total en filtre  $u$  pour les champs D2 et D3, qui est inférieur à ceux des autres champs.

Enfin le T0003, ajoute le temps de pose et revient à un point zéro de 30 pour tous les champs (il n’y a pas de table récapitulative de disponible au moment de l’écriture).

### 2.2.2 Photométrie - Comparaison avec le SDSS

La calibration photométrique est réalisée par Elixir, qui fournit les équations de transformation permettant de mesurer les magnitudes dans le système du



CFHTLS Field	CFHTLS Deep synoptic D1					CFHTLS Deep synoptic D2					CFHTLS Deep synoptic D3					CFHTLS Deep synoptic D4				
Filter	u	g	r	i	z	u	g	r	i	z	u	g	r	i	z	u	g	r	i	z
Stacked image (FITS. Selected images are listed in *.lst files)	Y	Y	Y	Y	Y	Y (CCD#03 missing)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
RA , DEC (J2000)	02:25:59 , -04:29:40					10:00:28 , +02:12:30					14:19:27 , +52:40:56					22:15:31 , -17:43:56				
Nb files stacked	16	37	48	100	34	7	24	17	38	28	7	37	64	127	42	28	56	78	118	-
Exp. time (s)	10560	7515	17280	52000	12240	4620	5400	6060	18500	10080	4620	8010	20820	59640	15120	16680	11250	26400	58800	-
Seeing (")	1.15	0.98	0.87	0.88	0.86	0.88	0.98	0.89	0.95	0.84	0.88	0.95	0.93	0.92	0.85	1.05	0.99	0.85	0.88	-
Mag. ZP	30.	30.	30.	30.	30.	30.	30.	30.	30.	30.	30.	30.	30.	30.	30.	30.	30.	30.	30.	-
Pixel size (")	0.1860	0.1860	0.1860	0.1860	0.1860	0.1860	0.1860	0.1860	0.1860	0.1860	0.1860	0.1860	0.1860	0.1860	0.1860	0.1860	0.1860	0.1860	0.1860	-
Image size (plx.)	20404 x 21547					20324 x 21442					19486 x 20482					20397x21475				
Effective FOV (=area in deg2)	0.80					0.69					0.77					0.77				
Chi2 image (FITS) [Y=included]	N	Y	Y	Y	N	N	Y	Y	Y	N	N	Y	Y	Y	N	N	Y	Y	Y	N
Dust Extinction Image (FITS, E(B-V))	Y					Y					Y					Y				
Weight map images (FITS)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
Completeness (plots)	png / ps					png / ps					png / ps					png / ps				
50% completeness limiting mag.	26.4	26.4	26.1	26.1	24.5	26.2	26.2	25.7	25.6	24.4	26.0	26.5	26.4	26.2	24.6	26.2	26.2	25.9	25.7	-

FIG. 2.6 – Résumé de l'état du T0001. Issue du site de térapix

Sloan. La définition des magnitudes est comme suit :

$$M_{instru}(FIL) = -2.5\log(DN) + 2.5\log(T) + PHOT_C + PHOT_K \cdot (AIRMAS - 1) \quad (2.11)$$

$$M_{SDSS}(FIL) = M_{instru} + PHOT_X \cdot (M_{SDSS}(PHOT_{C1}) - M_{SDSS}(PHOT_{C2})) \quad (2.12)$$

où FIL est le filtre utilisé pour obtenir ces données, DN est le nombre d'unités analogiques à numériques (ADU Analog to Digital Unit) mesuré dans l'image FITS, T est le temps d'exposition,  $PHOT_C$  le point zéro photométrique exprimé en ADU et mesuré par Elixir pour ce filtre,  $PHOT_K$  le terme de masse d'air pour ce filtre issu de la base de données pour Megacam, AIRMASS est la masse d'air au moment où l'image a été prise,  $PHOT_X$  est le terme de couleur issu de la base de données de Megacam associé aux filtres définis par les mots clés  $PHOT_{C1}$  et  $PHOT_{C2}$ .  $M_{SDSS}(PHOT_{C1})$  et  $M_{SDSS}(PHOT_{C2})$  sont les magnitudes dans le système SDSS dans les filtres  $PHOT_{C1}$  et  $PHOT_{C2}$ . D'une observation à l'autre la seule donnée mise à jour est la valeur du point zéro  $PHOT_C$ . Celle-ci n'est vraiment utilisable que lorsque les données sont prises dans des conditions photométriques, ce qui n'est pas toujours le cas. Il est donc nécessaire dans les traitements ultérieurs effectués par Térapix de corriger ce point zéro, et ce avant l'addition des données. Ces corrections sont calculées par l'outil SCAMP (celui-ci permet de lire les catalogues issus de SourceExtractor et de calculer

CFHTLS Target			Stack properties																								Effective Area	CFHTLS Ident			
Ident	RA	DEC	Magnitude ZP					Nb of files*					Total Exposure Time (seconds)					Seeing (arc-second)**					Completeness limit (80%/50%)***					(unmasked)	Ident		
W'(n,m) ref. pos.	J.2000	J.2000	u'	g	r	i	z	u'	g	r	i	z	u'	g	r	i	z	u'	g	r	i	z	u'	g	r	i	z	deg <sup>2</sup>	W'(n,m) ref. pos.		
W1-1+3	02:14:10	-04:12:00	29.92	29.97	29.93	30.00	30.00	4	4	2	7	10	2400.4	2000.4	1000.2	4305.7	6001.0	0.89	0.85	0.91	0.88	0.744	-25.5	-25.6	24.6/24.9	-24.9	-24.0	0.659663	W1-1+3		
W1-1+2	02:14:10	-05:08:00	30.03	30.04	29.97	30.00	29.98	5	8	6	7	11	3000.5	4000.8	3000.5	4305.7	6601.2	1.08	1.00	0.92	0.71	0.811	-25.3	-25.6	-25.0	-25.1	-24.2	0.629335	W1-1+2		
W1-0+3	02:18:00	-04:12:00	29.93	29.92	29.94	30.00	29.98	5	4	2	6	9	3000.5	2000.4	1000.2	3690.6	5400.7	0.97	0.95	0.96	0.88	0.988	-25.6	-25.6	24.5/24.8	-24.9	-23.6	0.569021	W1-0+3		
W1-0+2	02:18:00	-05:08:00	29.99	29.88	29.89	30.00	29.95	2	4	2	7	8	1200.3	2000.2	1000.2	4305.7	4800.7	1.00	0.94	0.94	0.65	0.933	24.9/25.2	-25.6	24.6/24.9	-25.4	-23.6	0.631755	W1-0+2		
W1-0+1	02:18:00	-06:04:00	30.07	29.97	29.96	30.00	29.95	3	5	3	7	11	1800.3	2500.5	1500.3	4305.7	6601.2	1.10	1.02	0.87	0.72	0.877	24.9/25.2	25.3/25.6	24.8/25.1	24.9/25.2	23.7/24.1	0.633995	W1-0+1		
W1+1+3*	02:21:50	-04:12:00	30.00	30.00	-	30.00	30.00	3	5	-	8	9	1800.3	2500.4	-	4920.9	5400.9	0.95	0.93	-	0.89	0.766	25.1/25.4	25.4/25.7	-	24.5/24.9	23.6/24.0	0.791551	W1+1+3*		
W1+1+2	02:21:50	-05:08:00	29.90	29.99	29.93	30.00	30.03	1	3	2	8	10	600.1	1500.2	1000.2	4920.8	6001.0	1.15	0.90	0.86	0.90	0.766	24.3/24.7	25.3/25.6	24.7/25.0	24.4/24.7	23.7/24.1	0.625136	W1+1+2		
W1+1+1	02:21:50	-06:04:00	29.94	29.93	29.96	30.00	30.05	1	6	2	7	13	600.1	3000.5	1000.2	4305.7	7801.3	0.98	0.96	0.96	0.96	0.877	24.6/24.9	25.4/25.7	24.6/24.9	24.6/24.9	23.8/24.2	0.660033	W1+1+1		
W1+2+3	02:25:39	-04:12:00	29.86	29.88	29.96	30.00	30.07	7	5	4	7	9	5950.3	2500.0	2000.4	4340.5	7200.4	1.12	0.93	0.78	0.87	0.677	25.3/25.6	25.4/25.7	25.0/25.3	-24.8	24.1/24.4	0.497407	W1+2+3		
W1+2+2	02:25:39	-05:08:00	30.00	29.96	29.96	30.00	29.99	2	4	2	7	10	1200.2	2000.4	1000.2	4340.7	6000.9	1.04	0.99	1.06	0.98	0.911	24.7/25.0	25.2/25.6	24.4/24.7	-24.8	-24.0	0.552322	W1+2+2		
W1+2+1	02:25:39	-06:04:00	30.05	29.96	29.99	30.00	30.05	5	2	2	8	10	3000.5	1000.2	1000.2	4960.7	6000.8	1.20	1.02	1.02	0.98	0.766	25.0/25.3	24.8/25.2	24.3/24.6	24.3/24.7	23.7/24.2	0.559442	W1+2+1		
W1+2+0	02:25:39	-07:00:00	-	30.00	30.00	-	-	-	4	2	-	-	-	2000.3	1000.2	-	-	-	0.95	1.02	-	-	-	25.3/25.6	24.3/24.6	-	-	-	0.663033	W1+2+0	
W1+3+3	02:29:29	-04:12:00	-	30.00	-	-	-	-	4	2	6	-	-	2000.2	-	-	-	-	0.99	-	-	-	-	25.2/25.5	-	-	-	-	0.484828	W1+3+3	
W1+3+2	02:29:29	-05:08:00	-	29.95	29.96	30.00	-	-	5	2	7	-	-	2500.4	1000.2	4340.3	-	-	0.95	0.89	0.79	-	-	25.3/25.6	24.5/24.8	-25.0	-	-	0.568010	W1+3+2	
W1+3+1	02:29:29	-06:04:00	-	29.94	29.95	30.00	-	-	5	2	6	-	-	2500.2	1000.1	3720.3	-	-	1.07	0.96	0.92	-	-	25.2/25.5	24.4/24.7	24.4/24.7	-	-	0.488128	W1+3+1	
W1+3+0	02:29:29	-07:00:00	-	29.90	29.93	30.00	-	-	5	4	6	-	-	2500.4	2000.3	3720.3	-	-	1.01	0.88	0.79	-	-	25.3/25.6	24.7/25.0	-24.9	-	-	0.468449	W1+3+0	
W1+4+3	02:33:19	-04:12:00	-	-	-	30.00	-	-	7	1	7	-	-	-	-	4340.3	-	-	-	-	1.00	-	-	-	-	24.3/24.6	-	-	-	0.482798	W1+4+3
W1+4+2	02:33:19	-05:08:00	-	29.96	29.99	30.00	-	-	10	2	6	-	-	5000.9	1000.2	3720.6	-	-	1.00	0.85	0.99	-	-	25.3/25.7	24.4/24.8	24.0/24.3	-	-	0.563741	W1+4+2	
W1+4+1	02:33:19	-06:04:00	-	30.00	29.99	30.00	-	-	4	2	7	-	-	2000.3	1000.2	4340.6	-	-	0.93	0.88	0.92	-	-	25.1/25.5	24.3/24.7	24.3/24.7	-	-	0.516895	W1+4+1	
W1+4+0	02:33:19	-07:00:00	-	29.97	29.98	30.00	-	-	5	2	8	-	-	2500.4	1000.2	4960.3	-	-	0.99	0.81	0.76	-	-	25.1/25.5	24.5/24.9	24.3/24.6	-	-	0.475189	W1+4+0	
W2-1+2	08:50:11	-01:27:00	-	30.00	30.00	-	-	-	5	2	-	-	-	2500.4	1000.2	-	-	-	0.84	1.02	-	-	-	25.3/25.8	24.3/24.8	-	-	-	0.651783	W2-1+2	
W2-0+3	08:54:00	-01:27:00	-	-	30.00	-	-	-	1	-	-	-	-	500.1	-	-	-	-	1.07	-	-	-	-	23.9/24.3	-	-	-	-	0.698519	W2-0+3	
W2+2+3	09:01:37	-01:27:00	-	30.00	30.00	-	-	-	5	1	-	-	-	2500.4	500.1	-	-	-	0.75	1.10	-	-	-	23.9/25.0	22.0/23.4	-	-	-	0.662633	W2+2+3	
W2+2+2	09:01:37	-02:23:00	-	-	30.00	-	-	-	4	-	-	-	-	2000.3	-	-	-	-	0.71	-	-	-	-	25.4/25.7	-	-	-	-	0.519135	W2+2+2	
W2+3+3	09:05:26	-01:27:00	-	29.96	30.00	30.00	-	-	2	3	2	-	-	1000.2	1500.2	1240.2	-	-	1.06	0.92	1.08	-	-	25.0/25.3	24.6/24.9	23.5/23.8	-	-	0.621536	W2+3+3	
W2+3+2	09:05:26	-02:23:00	-	29.91	29.97	30.00	-	-	4	2	4	-	-	2000.4	1000.2	2480.5	-	-	1.07	0.97	0.82	-	-	24.9/25.4	24.4/24.7	24.3/24.6	-	-	0.609247	W2+3+2	
W2+3+1	09:05:26	-03:19:00	-	30.00	30.00	-	-	-	7	2	-	-	-	3500.6	1000.2	-	-	-	1.02	0.91	-	-	-	25.2/25.6	24.5/24.9	-	-	-	0.641694	W2+3+1	
W2+3+0	09:05:26	-04:15:00	-	30.00	30.00	-	-	-	4	1	-	-	-	2000.4	500.1	-	-	-	0.93	1.02	-	-	-	25.0/25.4	23.9/24.3	-	-	-	-	W2+3+0	
W3-3+1	13:59:33	+53:34:31	-	29.99	30.02	30.00	-	-	4	2	6	-	-	2000.6	1000.3	3720.9	-	-	0.98	0.83	0.91	-	-	24.9/25.2	24.5/24.8	24.3/24.7	-	-	0.717048	W3-3+1	
W3-1+1	14:11:47	+53:34:31	-	30.00	-	-	-	-	3	-	-	-	-	660.6	-	-	-	-	0.95	-	-	-	-	24.2/24.5	-	-	-	-	0.837467	W3-1+1	
W3-0+2	14:17:54	+52:38:31	-	29.97	29.95	30.00	-	-	6	2	7	-	-	2500.8	1000.4	4341.4	-	-	0.93	0.94	0.76	-	-	25.2/25.6	24.4/24.8	-24.8	-	-	0.716148	W3-0+2	
W3-0+1	14:17:54	+53:34:31	-	-	-	30.00	-	-	-	-	6	-	-	-	-	3721.2	-	-	-	0.80	-	-	-	-	24.4/24.7	-	-	-	-	0.698999	W3-0+1
W3-0+0	14:17:54	+54:30:31	-	30.00	-	-	-	-	3	-	-	-	-	1500.5	-	-	-	-	1.04	-	-	-	-	25.1/25.4	-	-	-	-	0.732626	W3-0+0	
W3+1+3	14:23:47	+51:42:31	-	30.03	30.00	30.00	-	-	4	2	6	-	-	2001.0	1000.5	3721.6	-	-	0.95	0.86	0.83	-	-	25.3/25.6	24.8/25.1	24.7/25.0	-	-	0.644664	W3+1+3	
W3+1+2	14:23:54	+52:38:31	-	30.00	-	30.00	-	-	5	-	7	-	-	2501.3	-	4341.7	-	-	0.92	-	0.87	-	-	25.4/25.7	-	24.3/24.7	-	-	0.746335	W3+1+2	
W3+1+1	14:24:01	+53:34:31	-	30.00	29.98	30.00	-	-	5	2	7	-	-	2501.1	1000.4	4342.0	-	-	1.00	0.89	0.80	-	-	25.5/25.7	24.8/25.1	24.5/24.9	-	-	0.752785	W3+1+1	
W3+1+0	14:24:09	+54:30:31	-	30.00	30.00	30.00	-	-	5	1	7	-	-	2501.2	500.2	4341.9	-	-	0.96	0.83	0.84	-	-	25.4/25.7	24.5/24.8	24.6/24.8	-	-	0.805060	W3+1+0	
W3+2+3	14:29:39	+51:42:31	-	29.90	29.96	30.00	-	-	6	2	10	-	-	3001.4	1000.5	6202.4	-	-	1.02	0.68	0.74	-	-	25.3/25.5	24.9/25.1	24.8/25.2	-	-	0.798671	W3+2+3	
W3+2+2	14:29:53	+52:38:31	-	29.96	29.96	30.00	-	-	7	3	7	-	-	3501.4	1500.6	4341.7	-	-	0.98	0.94	0.72	-	-	-25.7	24.5/24.8	-25.1	-	-	0.754645	W3+2+2	
W3+2+1	14:30:08	+53:34:31	-	29.96	29.95	30.00	-	-	3	1	10	-	-	1500.7	500.2	6202.2	-	-	1.01	1.07	0.83	-	-	25.2/25.5	24.1/24.5	24.6/24.9	-	-	0.711638	W3+2+1	
W3+2+0	14:30:23	+54:30:31	-	30.00	30.00	-	-	-	1	1	-	-	-	500.2	500.2	-	-	-	1.00	1.01	-	-	-	24.6/25.0	24.2/24.6	-	-	-	0.784452	W3+2+0	
D1	02:25:59	-04:29:40	30.02	29.97	29.99	30.00	30.04	54	100	161	239	150	35645.8	19310.8	52511.7	118860.0	54011.9	1.06	0.96	0.93	0.91	0.87	26.0/26.4	25.8/26.3	25.4/26.1	25.1/25.9	24.4/24.9	0.800	D1		
D2	10:00:28	+02:12:30	30.00	30.00	30.00	30.00	30.00	7	29	16	73	30	4620.7	6527.3	15424.0	35287.8	10803.0	0.89	0.99	0.90	0										

les solutions astrométriques et photométriques pour n'importe quelle séquence d'images) en minimisant la somme quadratique et pondérée des différences de magnitude entre plusieurs détections d'un même objet. Les magnitudes mesurées via une détermination automatique de la taille occupée par l'objet sur l'image, notées *MAG\_AUTO* dans SourceExtractor ont été utilisées car moins sensibles aux variations de seeing (ceci dépend tout de même de la densité d'objet dans le champ étudié). En effet, ces magnitudes sont issues de l'algorithme de Kron (1980) : dans un premier temps une ellipse est définie autour de l'objet, puis dans cette ellipse est calculé le "premier moment" :

$$r_1 = \frac{\sum rI(r)}{\sum I(r)} \quad (2.13)$$

avec  $I$  l'intensité du flux à la distance  $r$  du centre.

Il est important de noter qu'il n'y a pas de champ de référence pour la calibration photométrique, mais comme certains champs ont été pris dans des conditions photométriques, ils peuvent être utilisés comme référence pour le point zéro. Les erreurs dues à l'utilisation d'une telle méthode sont de l'ordre de 1%-2%. Ainsi lors de la détermination du point zéro pour une image, Térapix a décidé de relâcher les contraintes sur les valeurs de point zéro déterminées par Elixir afin de permettre à ces valeurs de varier quelque peu. Ceci permet d'estimer les erreurs qui sont faites sur la détermination de ces points zéro ; pour cela il faut considérer des objets à très haut signal sur bruit (supérieur à 100), sur ceux-ci les variations de magnitude dues à la variation du point zéro sont de l'ordre de 0.02-0.05 magnitudes. De plus, pour toutes les images du 'Wide' et du 'Deep' les points zéro ont été réajustés (de l'ordre de 0.05 magnitudes) à partir des diagrammes couleur-couleur. Il a été choisi un champ de référence, le D2 (de façon arbitraire), et les points zéro ont été recalés pour que les étoiles dans les diagrammes couleur occupent toutes le même locus. Néanmoins, on peut s'interroger sur le choix du champ de référence puisque parmi les champs du 'Deep', D2 est celui qui possède le moins de temps d'exposition tous filtres confondus.

Les catalogues fournis par le CFHT sont faits en utilisant l'outil SourceExtractor (Bertin et Arnouts). Ceux-ci contiennent notamment les magnitudes des objets détectés. Je renvoie au manuel du programme pour connaître les détails de ce qui est fourni dans les catalogues, mais je vais rappeler certains points qui me paraissent essentiels. Tout d'abord le catalogue fournit les magnitudes de type *MAG\_AUTO*, calculées selon le système AB, qui permettent de limiter les écarts entre vraie magnitude et magnitude calculée (grâce à la méthode 'automatic aperture'). De plus les erreurs photométriques associées à ces magnitudes sont calculées à partir : i) de la surface de détection de l'objet, ii) de son flux en ADU, iii) de l'écart type du bruit en ADU venant de l'estimation du ciel et iv) des caractéristiques du détecteur. Mais, dicit le manuel :

**"This error must be considered only as a lower value since it does not take into account the complex uncertainty on the local background estimate."** De plus il convient de tenir compte des erreurs de calibration photométrique citées dans le paragraphe précédent. SourceEXtractor permet aussi d'attribuer des drapeaux de qualité photométrique sur chaque détection.

Les filtres utilisés pour le CFHTLS sont peu différents de ceux du SDSS (cf. figure 2.8), si ce n'est pour quelques points : tout d'abord les filtres  $i$  et  $z$  de

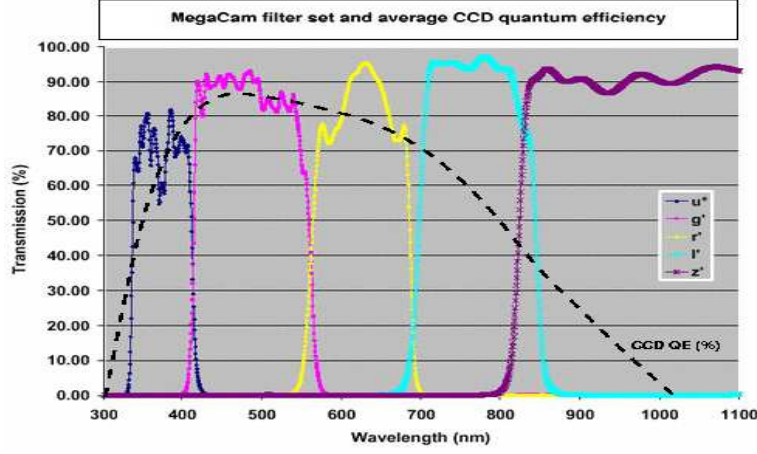


FIG. 2.8 – Transmission des filtres du CFHT et efficacité quantique moyenne du détecteur.

MegaCam se superpose sur un petit domaine, ce qui a pour effet général par rapport aux couleurs du SDSS de diminuer la couleur  $i-z$ . Ceci a surtout un effet pour la sélection de quasars à grand décalage spectral qui se fait essentiellement grâce à cette couleur (Fan et al. 1999, Willott et al. 2006). De plus le filtre  $u$  est plus sensible dans l'extrême bleu du spectre visible.

De la même façon que l'équipe de Térapix a recalibré les points zéro à partir de diagrammes couleur-couleur en utilisant des simulations d'étoiles avec les spectres de Pickles 1998 (ces diagrammes peuvent être retrouvés sur le site de Térapix), j'ai voulu comparer mes simulations d'étoiles (décrites avec l'outil HyperZ dans la section 4.1 : elles sont obtenues en simulant les couleurs des étoiles via des spectres d'étoiles pris de façon aléatoire parmi l'échantillon de Pickles) réalisées à travers les filtres SDSS et CFHT, aux données. Dans un premier temps j'ai effectué cette comparaison par rapport aux données CFHTLS du D3 pour  $i < 22$ , puis dans un deuxième temps aux données du SDSS. J'ai aussi voulu comparer les observations du SDSS aux observations du CFHTLS du D3 pour  $i < 22$ . Le tout revient à comparer les effets dus aux différences entre les filtres et les systèmes optiques (en supposant que la photométrie suit un traitement équivalent).

Les diagrammes comparant CFHTLS et simulations (cf. figure 2.9) sont en bon accord, les locus des étoiles se retrouvent globalement dans les mêmes régions, avec bien sûr une dispersion plus importante pour les données. On peut éventuellement voir dans le plan  $r-i/g-r$  que la couleur  $r-i$  aurait tendance à être un peu supérieure aux simulations, mais cela ne se retrouve pas dans le plan  $i-z/r-i$ . La figure 2.10 représente la même chose mais pour  $i < 20$  et les figures sont issues du site de Térapix, le nombre d'objets est moins important que pour  $i < 22$ , et le décalage est moins visible. Les diagrammes comparant observations du SDSS et simulations (cf. figure 2.11) présentent un parfait accord, le léger décalage observé dans le plan  $r-i/g-r$  semble plus faible que pour les données



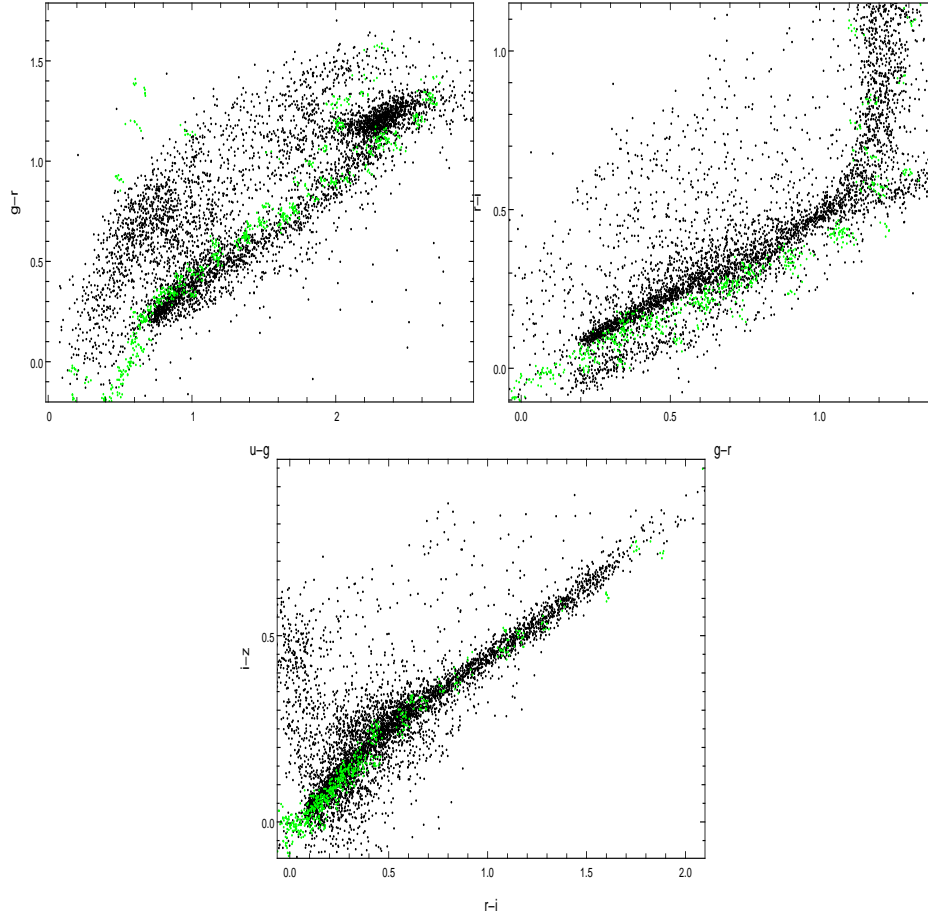


FIG. 2.9 – Comparaison dans les diagrammes couleur couleur des simulations des étoiles selon Pickles 1998 (en vert) et des observations du CFHTLS D3 pour  $i < 22$  (en noir).

du CFHTLS.

En ce qui concerne les comparaisons SDSS et CFHTLS (cf. figure 2.12), les étoiles du plan  $u - g/g - r$  se superposent bien, tandis que pour les autres plans un décalage et une légère déformation du locus se retrouve. Ceci est dû essentiellement aux différences entre les filtres.

Il est possible de déterminer les différences entre les systèmes de filtres, par exemple en calculant les magnitudes dans chacun des systèmes pour une étoile standard comme Vega (exemple de référence Fukugita et al. 1995). Le calcul pour les filtres du SDSS a déjà été réalisé, j'ai néanmoins recalculé ces valeurs avec les données des filtres que je vais utiliser.

	AB magnitude de Vega				
	$u$	$g$	$r$	$i$	$z$
CFHT	0.416	-0.069	0.154	0.401	0.54
SDSS	0.955	-0.094	0.157	0.393	0.538

Pour se rendre compte de la difficulté de convertir un système photométrique

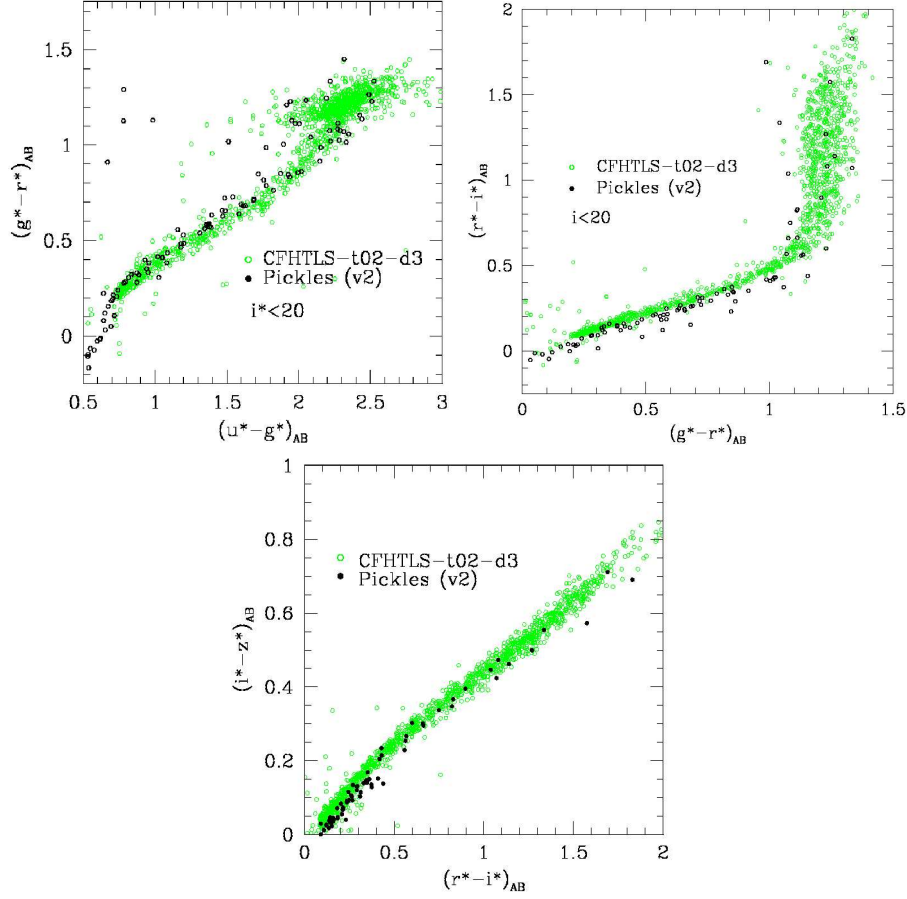


FIG. 2.10 – Comparaison dans les diagrammes couleur-couleur des simulations des étoiles selon Pickles (en noir) et des observations du CFHTLS D3 pour  $i < 20$  (en vert). Issu du site de Terapix : [http://terapix.iap.fr/rubrique.php?id\\_rubrique=198](http://terapix.iap.fr/rubrique.php?id_rubrique=198)

en un autre je renvoie à Jester et al. (2005) qui réalisent la conversion du système SDSS au système standard UBVRI. Ils réalisent cette conversion en choisissant et en adaptant les relations de corrélation entre les couleurs des deux systèmes. Pour faire la même chose avec le système de filtres de Megacam, j'ai tracé les couleurs  $u - g$  en système SDSS en fonction de  $u - g$  en système CFHT (cf. figure 2.13) pour des quasars de  $z$  de 0 à 7 en bleu et de 0 à 2 en rouge. Ces couleurs sont fortement corrélées, mais l'erreur que l'on commettrait en adaptant par exemple un polynôme à cette relation de corrélation serait supérieur aux erreurs photométriques des relevés. Faire une conversion dans ces systèmes pour tous les quasars ne me semble pas approprié puisque les erreurs sur la relation ainsi trouvée ne nous permettraient pas de l'utiliser pour la sélection avec les objets issus du CFHTLS.

Au final, il est important de noter que les points communs entre le SDSS et le CFHTLS sur le plan photométrique permettent de profiter de l'acquis du SDSS. Par exemple, la recherche de quasars à très grand décalage spectral,

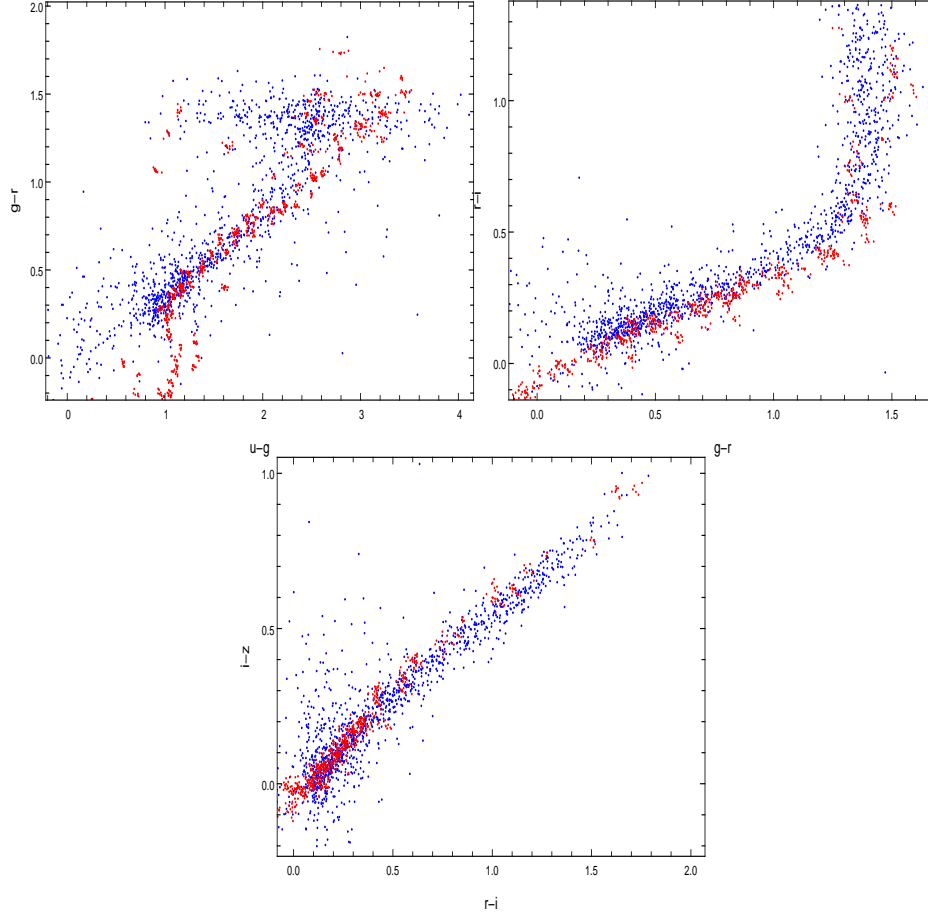


FIG. 2.11 – Comparaison dans les diagrammes couleur couleur des simulations des étoiles selon Pickles 1998 (en rouge) et des observations du SDSS D3 pour  $i < 22$  (en bleu).

comme présenté par Fan et al. (1999) a été reprise par Wilott et al. (2006). Les modifications apportées par Wilott sont assez minimes, et tiennent compte des différences présentées entre les filtres.

### 2.3 Autres champs observés au CFHT

Nous avons aussi eu accès à des données privées ayant subi le même traitement que celles du relevé CFHTLS. Ces champs ont la particularité de posséder un quasar brillant de grand décalage spectral au centre. De plus, pour ce quasar, un spectre (obtenu grâce au spectrographe UVES) de très haute qualité est déjà disponible. Les profondeurs en magnitudes pour ces champs sont de 25.5, 26.5, 25.7, 25.5 et 24.0 pour les filtres  $u, g, r, i, z$  respectivement. Les champs observés sont résumés dans la table 2.2.

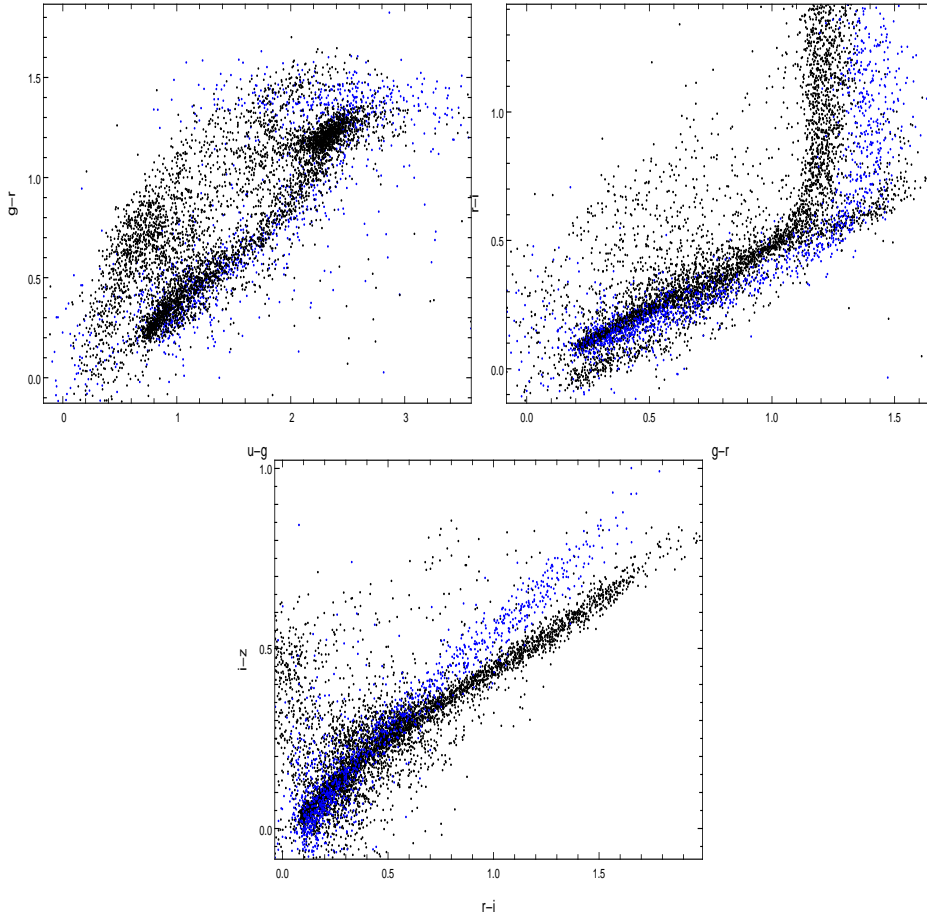


FIG. 2.12 – Comparaison dans les diagrammes couleur couleur des observations du SDSS D3 pour  $i < 22$  (en bleu) et des observations du CFHTLS D3 pour  $i < 22$  (en noir).

Objets au centre du champ	RA	DEC
HE0940-1050	09 :42 :53.6	-11 :04 :27.0
Q1232+082	12 :34 :37.5	+07 :58 :40.5

TAB. 2.2 – Champs observés au CFHT avec les quasars situés en leur centre ainsi que leur coordonnées.



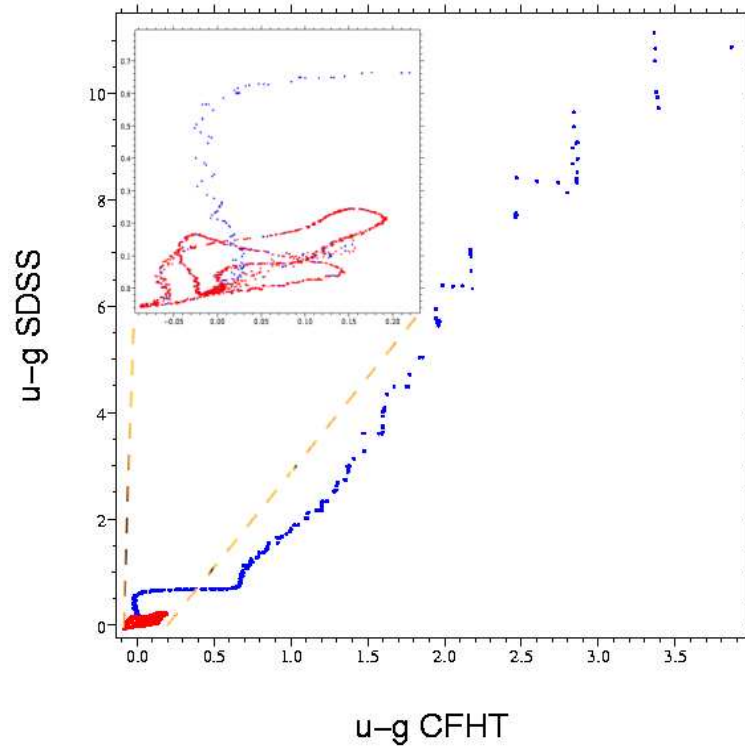


FIG. 2.13 –  $u - g$  SDSS en fonction de  $u - g$  CFHT pour des quasars simulés de décalage spectral entre 0 et 7 en bleu et de 0 à 2 en rouge

## Chapitre 3

# Présentation des méthodes de sélection de quasars

*"There's an inverse relationship between how good something is for you and how much fun it is."*

– Calvin

## Sommaire

---

<b>3.1</b>	<b>Les propriétés des quasars . . . . .</b>	<b>61</b>
3.1.1	Spectre en émission des quasars . . . . .	61
	Domaine radio . . . . .	62
	Infra-rouge . . . . .	63
	Domaine X . . . . .	63
	Optique et ultra-violet . . . . .	64
3.1.2	Les absorptions dans les spectres des quasars . . . . .	66
	Systèmes métalliques . . . . .	66
	Systèmes de la forêt Lyman- $\alpha$ . . . . .	67
	Larges raies d'absorption (BAL) . . . . .	67
<b>3.2</b>	<b>Introduction à la sélection de quasars . . . . .</b>	<b>69</b>
<b>3.3</b>	<b>Diagrammes couleur-couleur . . . . .</b>	<b>71</b>
3.3.1	Le locus des étoiles dans les diagrammes couleur-couleur . . . . .	72
3.3.2	Position des galaxies dans les diagrammes couleur-couleur . . . . .	74
3.3.3	Position des quasars dans les diagrammes couleur-couleur . . . . .	75
3.3.4	Algorithmes de sélection des quasars dans le SDSS . . . . .	79
	Premier algorithme de sélection du SDSS . . . . .	79
	Seconde sélection du SDSS . . . . .	86
	Coupe dans les diagrammes couleur-couleur . . . . .	88
3.3.5	Contaminants possibles à ce type de sélection . . . . .	89
<b>3.4</b>	<b>Adaptation de distribution spectrale d'énergie . . . . .</b>	<b>91</b>
3.4.1	Historique et principe . . . . .	92
3.4.2	Comparaison aux diagrammes couleur-couleur . . . . .	96

---

Le principe même de l'identification d'objets, les quasars pour ce qui nous concerne, est de comparer une base de connaissances à de nouvelles observations. La base de connaissances est ici un ensemble de spectres ou de distributions spectrale d'énergie (DSE) des quasars connus. Cet ensemble de DSE est censé regrouper l'ensemble des propriétés connues des quasars en excluant les cas extrêmes ou exotiques. Pour construire cet ensemble de DSE, il est nécessaire d'avoir un maximum d'informations sur les caractéristiques des quasars. Ce sont les propriétés spectrales des quasars qui importent puisqu'elles vont conditionner les magnitudes observées et ainsi leurs couleurs. Il convient de collecter un maximum de spectres de quasars aux propriétés différentes, telles que leur luminosité intrinsèque, leur décalage spectral, leur obscurcissement.... On peut en tirer une sorte de moyenne : Francis et al. (1991). Il est possible aussi d'en tirer plus qu'une simple moyenne, voire de classer ces quasars en différentes sous classes : Yip et al. (2004). La comparaison du spectre d'un objet (ou bien les couleurs induites par ce spectre) à une DSE observée particulière permettra de classer l'objet en quasar ou de l'identifier à d'autres classes d'objets, étoiles ou galaxies. La réalisation et la compréhension de ces DSE est donc particulièrement essentielle.

## 3.1 Les propriétés des quasars

### 3.1.1 Spectre en émission des quasars

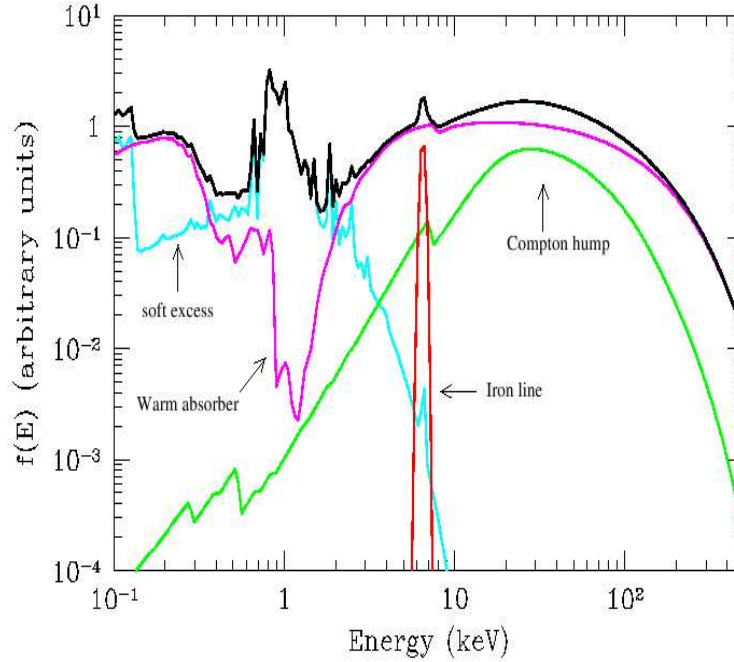


FIG. 3.1 – Le spectre moyen (en trait noir et épais) et les composantes principales (en couleur) d'un noyau actif de type I. "Supermassive black holes in the distant universe" Amy J. Barger.

Avant d'aller plus loin, je propose de définir un quasar : il s'agit d'un objet d'apparence ponctuelle parfois identifié comme une source radio. Le flux issu de son continu est variable, sur des jours ou des mois. Il émet largement dans l'ultra-violet comparé à la plupart des étoiles. Il montre de larges raies d'émission (plus de  $5000 \text{ km.s}^{-1}$  FWHM). C'est une source de rayons X, ce qui est sans doute une propriété clé, et enfin il est situé à un grand décalage spectral (au-delà de l'univers local).

D'une façon générale le terme "noyau actif de galaxie", ou NAG, implique l'existence de phénomènes énergétiques dans le noyau ou dans les régions centrales des galaxies, phénomènes qui ne peuvent pas être clairement et directement associés aux étoiles ou à la formation d'étoiles (e.g. par opposition aux galaxies présentant une activité de formation d'étoiles, dites de type 'starburst'). Les deux classes les plus répandues de NAG sont les galaxies de type Seyfert et les quasars, la différence fondamentale entre ces deux sous-classes provenant de la quantité d'énergie émise par le noyau. Dans le cas des Seyfert celle-ci est comparable à l'énergie émise par les étoiles présentes dans la galaxie, tandis que dans un quasar la source centrale est plus brillante que la galaxie hôte d'un facteur 100. A priori les phénomènes qui créent ce rayonnement n'entrent pas dans notre étude, je vais simplement rappeler ce qui a été dit dans l'introduction.

L'émission du noyau est probablement due à une accréation de matière autour d'un trou noir super-massif et à la présence de gaz fortement ionisé à proximité de la zone centrale. La différenciation entre les types de NAG est due à une différence de point de vue de la ligne de visée : la présence d'un tore de poussières encerclant la zone d'émission peut occulter certaines parties des radiations suivant l'angle de vue. D'autre part la différenciation peut être évolutive, avec des noyaux de galaxies plus ou moins actifs. En supposant que les mécanismes d'émission sont les mêmes pour tous les NAG, quasars compris, je parlerai indifféremment de noyau actif de galaxie ou de quasar pour caractériser les différents domaines d'énergie de leur émission.

L'émission d'énergie se répartit sur un intervalle très large, du domaine radio aux rayons gamma, et les quasars sont parmi les sources les plus brillantes du ciel. Leur spectre se caractérise par un continu qui peut être modélisé par une loi de puissance dans chacun des domaines d'énergie (cf. équation 3.1, qui ne se retrouve pas en infra-rouge) sur lequel se superposent des raies d'émission :

$$F_\nu = C \nu^{-\alpha} \quad (3.1)$$

Avec  $\alpha$  l'indice de la loi de puissance,  $C$  une constante et  $F_\nu$  le flux par unité de fréquence, généralement mesuré en  $\text{ergs s}^{-1} \text{cm}^{-2} \text{Hz}^{-1}$ . Je vais passer en revue les caractéristiques du spectre sur différents domaines d'énergie :

### Domaine radio

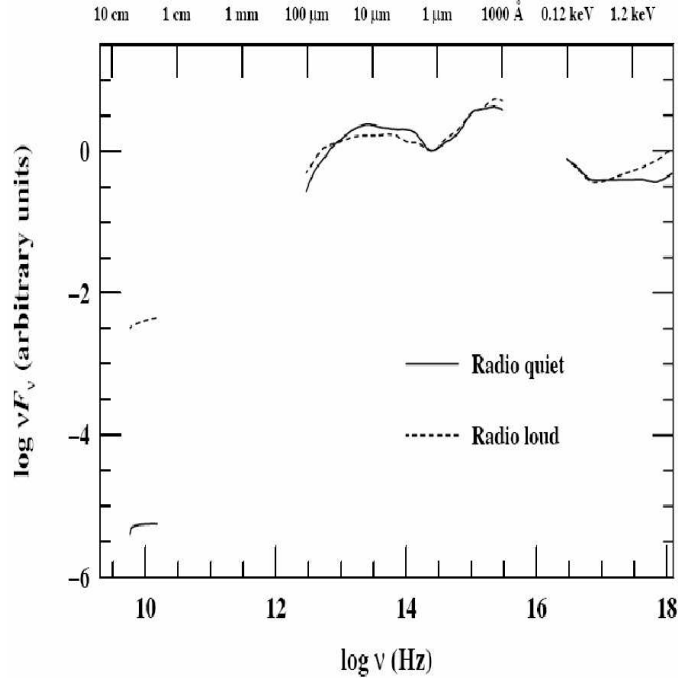


FIG. 3.2 – DSE comparant un échantillon de quasar 'radio-quiet' et 'radio-loud'. Issue de Sanders (1989).

Les objets présentant des émissions radio peuvent se séparer en deux catégories, l'une dominée par les lobes qui se situent autour de la source centrale, l'autre dominée par le centre. Dans le premier cas l'indice décrivant le spectre varie entre 0.5 et 1, les plus grands indices tendant à être à plus grand décalage spectral. Et dans le second cas, la plupart des sources ont un spectre quasiment plat. Les objets qualifiés de 'radio-loud' présentent un continu fort qui s'étend du domaine radio jusqu'à l'infra-rouge lointain (10% des quasars sont 'radio-loud'). Les objets qualifiés de 'radio-quiet' ont un spectre avec un indice  $\alpha > 2.5$  dans l'infra-rouge, qui semble indiquer la présence de poussières et l'émission radio est très faible (néanmoins détectable grâce à la sensibilité des détecteurs récents). L'origine de l'émission en radio est bien comprise, contrairement aux domaines optiques ou en rayons X. Le phénomène impliqué est un rayonnement synchrotron proposé par Alfen & Herlofson (1950) et Shklovsky (1953). La théorie peut être retrouvée dans Jackson (1975), Rybicki & Lightman (1979) et Krolik (1999).

### Infra-rouge

La partie infra-rouge du spectre n'est pas décrite par une loi de puissance unique. Le spectre des noyaux actifs de galaxies dans l'infra-rouge présente un minimum vers  $1 - 2\mu m$ , une bosse vers  $10 - 30\mu m$  due à l'émission thermique provenant de poussières et un déclin avec un indice  $\alpha > 3$  à des longueurs d'onde plus grandes. Le continu observé est généralement reproduit par plusieurs types d'activité : 'starburst' ou par une réémission de l'énergie provenant du quasar diffusé sur des poussières ; la question n'est pas tranchée (Elitzur et al. 2004). Récemment un composite dans ces longueurs d'onde a été réalisé (cf. figure 3.3) par Glikman et al. (2006). Ils proposent un modèle avec à la fois une loi de puissance d'indice  $\alpha = -0.92$  et une température de corps noir de 1260 K. Il faut néanmoins être prudent luminosité et décalage spectral sont souvent corrélés dans des échantillons limités en flux et donc les comparaisons entre domaine infra-rouge et optique ou ultra-violet le sont sûrement aussi.

### Domaine X

Comme on peut le voir sur la figure 3.1, le continu de l'émission en X peut être considéré en première approximation comme une loi de puissance qui s'étend de  $1keV$  jusqu'à plus de  $100keV$  sur laquelle se superposent des émissions et des absorptions. L'indice spectral typique est de l'ordre de  $\alpha \sim -0.8, \alpha \sim 1$ , jusqu'à des énergies de l'ordre de  $80 keV$  où un changement semble être observé par les données de BeppoSAX. En plus de cette émission on observe une composante de l'ordre de  $0.2$  à  $1keV$  dont l'origine n'est pas encore déterminée (émission en rayon X mou ou 'soft X-rays'). Une partie de l'émission doit être réfléchiée par du gaz ionisé, dont la principale caractéristique est un pic aux alentours des  $30keV$  et une importante décroissance vers  $4 - 5keV$ . Une caractéristique majeure est la présence d'une raie d'émission à  $6.4keV$  avec une largeur équivalente de l'ordre de  $100 - 200eV$ . Elle est due à la transition  $Fe - Kn = 2 - 1$  du fer, probablement due à un phénomène de fluorescence dans les parties internes du

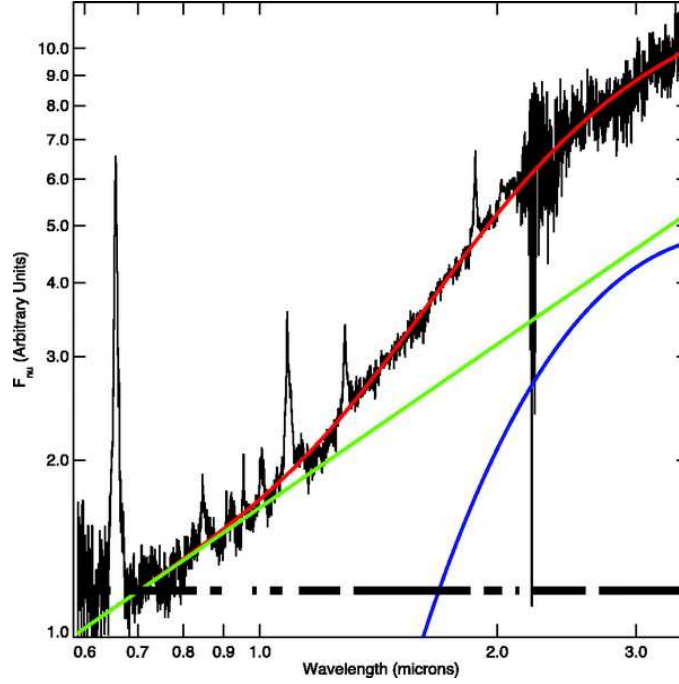


FIG. 3.3 – DSE de quasars dans l'infrarouge (en noir), montrant l'adaptation (en rouge) d'un modèle par une loi de puissance avec un indice spectral de -0.92 (en vert) et d'une courbe de corps noir à une température de 1260 K (en bleu). Figure issue de Glikman et al. (2006).

disque d'accrétion (Fabian et al. 2004).

### Optique et ultra-violet

L'émission dans ce domaine d'énergie est caractérisée par ce qui est appelé le 'Big Blue Bump' (Shields 1978, Malkan & Sargent 1982, Elvis 1985) où se situe le pic de l'émission des quasars, proche de la limite de Lyman à 1216 Å. Les différents relevés en optique (Francis et al. 1991, Croom et al. 2002, Brotherton et al. 2001, Vanden Berk et al. 2001) ont conduit à extraire quelques propriétés importantes :

De 300 à 5000 Å le continu peut être modélisé par deux lois de puissance : avec un indice de l'ordre de -1.8 entre 300 et 1200 Å et un indice de l'ordre de -0.4 entre 1200 et 5000 Å. Pour des longueurs d'onde plus grandes, le spectre s'aplatit. Des centaines de raies d'émission sont présentes dans les spectres, les plus fortes sont données dans la table 3.1. Les raies d'émission liées aux transitions permises sont larges (2000 - 15000  $km s^{-1}$ ) et celles liées aux transitions dites interdites sont étroites, une centaine de  $km s^{-1}$ . Un petit nombre de quasars (10-20 %) présentent des raies d'absorption très larges, parfois saturées avec des largeurs et des décalages vers le bleu de l'ordre de plusieurs milliers de  $km s^{-1}$ , jusqu'à 20 000 ou 30 000  $km s^{-1}$ , quasars qualifiés de BAL (Broad Absorption Line ; Weymann et al. 1991 ; Hamann et al. 1993).

En résumé l'émission du continu peut se séparer en plusieurs domaine :

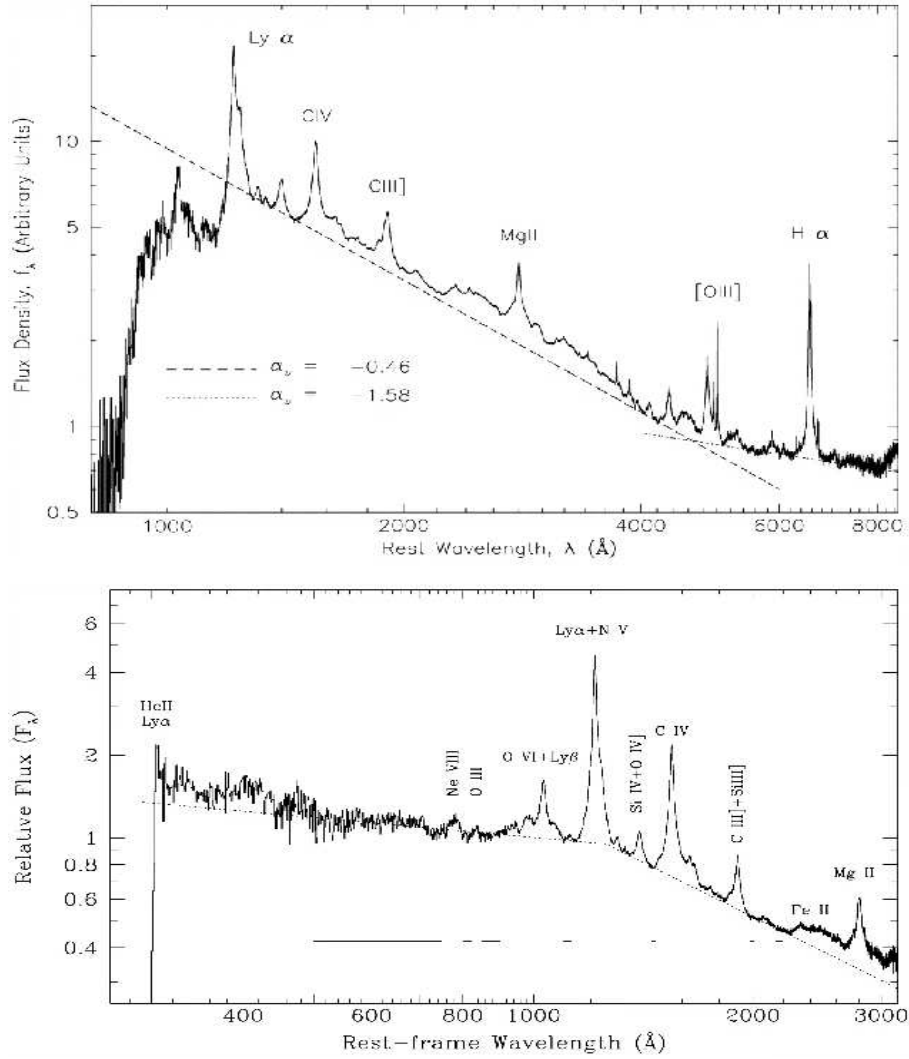


FIG. 3.4 – Panneau du haut, composite issu du SDSS Vanden berk et al. (2001), panneau du bas issu du HST Telfer et al. (2002). Les lignes sont les adaptations de loi de puissance. Issue de Risaliti et al. (2004).

1. Radio, avec une émission synchrotron des électrons relativistes, en générale négligeable par rapport aux émission dans d'autres longueurs d'ondes sauf pour les NAG fort émetteurs radio comme les quasars et les blazar.
2. mm-X, relativement plat mis à part deux bosses en infra-rouge et en ultra-violet. Pour le premier de  $2 - 5\mu m$ , l'émission de la poussière chauffé par l'émission centrale. Pour le second de 100 à 3000 Å, présence du 'Big Blue Bump' qui caractérise les NAG et est attribué à l'émission thermique de gaz chaud ( $10^5 K$ ) de la partie optiquement épaisse du disque d'accrétion autour du trou noir super massif central.



<i>Raie</i>	$\lambda(\text{\AA})$	$\langle \text{LE}() \rangle$	<i>Ref.</i>	<i>Raie</i>	$\lambda(\text{\AA})$	$\langle \text{LE}() \rangle$	<i>Ref.</i>
CIII+NIII	980	$9.7 \pm 0.2$	(1)	MgII	2799	$32.3 \pm 0.1$	(2)
OVI+Ly $\beta$	1030	$15.6 \pm 0.3$	(1)	H $\delta$	4103	$5.05 \pm 0.06$	(2)
Ly $\alpha$	1216	$91.8 \pm 0.7$	(1)	H $\gamma$	4341	$12.6 \pm 0.1$	(2)
N V	1240	$18.5 \pm 0.5$	(1)	H $\beta$	4863	$46.2 \pm 0.2$	(2)
SiIV+OIV]	1397	$8.13 \pm 0.09$	(2)	[OIII]	5007	$13.2 \pm 0.2$	(2)
CIV	1549	$23.8 \pm 0.1$	(2) <sup>a</sup>	H $\alpha$	6565	$194.5 \pm 0.$	(2)
CIII]	1909	$21.2 \pm 0.1$	(2)				

TAB. 3.1 – Raies d’émission avec Largeur Équivalente (LE)  $> 5 \text{ \AA}$  dans les spectres optique et ultra-violet des quasars. (1) Telfer et al. (2002) (Pour les raies  $\lambda < 1300 \text{ \AA}$ ); (2) Vanden Berk et al. (2001) (pour  $\lambda > 1300 \text{ \AA}$ ).

### 3.1.2 Les absorptions dans les spectres des quasars

Les spectres de quasars sont marqués par la présence de raies d’absorption. Celles-ci peuvent être classifiées en trois groupes selon la nature de ce qui cause cette absorption. Tout d’abord celles qualifiées d’intrinsèques qui sont dues à du gaz se trouvant aux alentours du noyau actif, qui ont lieu dans des régions très proches de celui-ci. Ensuite les absorptions produites par du gaz qui n’est pas lié au noyau mais à sa galaxie hôte. Et enfin les absorptions dues au gaz interposé le long de la ligne de visée qui n’est pas physiquement relié au noyau ou à sa galaxie hôte.

Entre l’observateur et le quasar il est possible que le long de la ligne de visée se trouvent des objets, tels que des galaxies, dont la présence peut être décelée dans le spectre des quasars par l’observation de raies en absorption à des décalages spectraux  $z_{\text{absorption}} < z_{\text{émission quasar}}$ . Ainsi chaque quasar peut présenter des centaines de raies d’absorption dans son spectre, ceci dépendant du décalage spectral où se situe le quasar et du domaine d’énergie observé. Les raies d’absorption les plus communément observées sont celles liées aux transitions Ly- $\alpha$  1216, C IV 1548,1551, et Mg II 2795,2802. Les systèmes ont été classifiés (Weymann et al. 1981, Sargent et al. 1988) :

#### Systemes métalliques

Ces raies nécessitent des observations à très haute résolution pour être résolues et les largeurs des différentes composantes sont généralement de l’ordre d’une dizaine de  $km s^{-1}$ , ce qui semble indiquer que les absorptions sont dues à différents nuages proches dans l’espace des vitesses. Les densités de colonne en hydrogène de ces systèmes sont de l’ordre de  $10^{17} - 10^{21} cm^{-2}$ . Une sous-classe très importante dans ces systèmes sont les systèmes lorentziens Lyman- $\alpha$  (SLLA), dont l’absorption est très forte et large (Wolfe et al. 1986) et dont les densités de colonne en hydrogène sont supérieures à  $2.10^{20} \text{ atomes } cm^{-2}$ . Il semblerait que ces systèmes soit associés à des halos de galaxies (Teplitz et al. 1998), et les systèmes d’absorption ne se limitent pas à l’hydrogène mais présentent bien de nombreuses absorptions métalliques. Il est possible d’observer des absorptions moléculaires dans ces systèmes, telles que  $H_2$ , (Levshakov

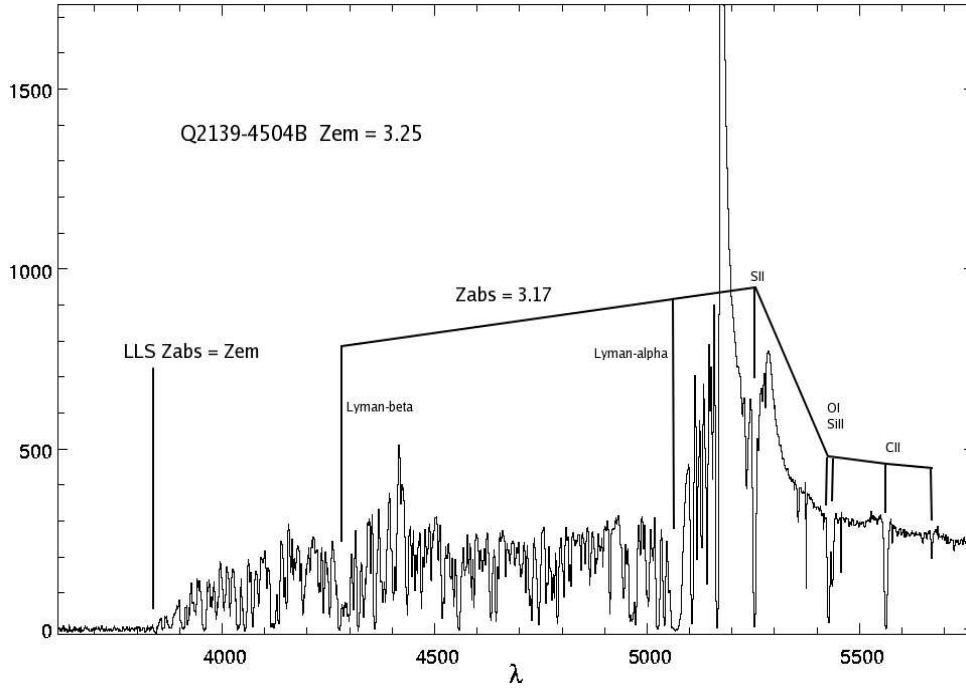


FIG. 3.5 – Quasar Q2139-4504B, avec le système lorentzien à  $z_{abs} = 3.17$  et les systèmes métalliques associés ainsi qu'un système LLS au même décalage spectral que l'émission du quasar.

& Varshalovich 1985; Foltz et al. 1988; Srianand, Petitjean & Ledoux 2000; Ledoux, Petitjean & Srianand 2003, Ledoux et al. 2006), où il pourrait aussi y avoir des preuves de la présence de poussières. Au titre de système d'absorption particulier, on peut citer les systèmes où la limite de Lyman est détectée en absorption, c'est-à-dire optiquement épais à  $912 \text{ \AA}$ , appelés 'Lyman limit systems' (LLS, Tytler 1982).

### Systèmes de la forêt Lyman- $\alpha$

Dans la partie bleue du spectre en-deçà de l'émission Lyman- $\alpha$  du quasar, le nombre de raies d'absorption par unité de longueur d'onde (densité de systèmes) augmente énormément. Ce phénomène a été observé par Lynds (1971) et expliqué en partie par Sargent et al. (1980). La plupart de ces raies sont attribuées à des absorptions Lyman- $\alpha$  de HI, et cette partie du spectre s'est vue attribué le nom de forêt Lyman- $\alpha$  (Weymann et al. 1981). La présence de métaux dans ces systèmes, jusqu'à de faibles densités de colonne ( $10^{13} \sim 10^{14} \text{ cm}^{-2}$ ), est détectée mais toujours avec des métallicités très faibles, inférieures au centième solaire (Tytler & Fan 1994).

### Larges raies d'absorption (BAL)

Dans les spectres de certains quasars, entre 10% et 30% (Tolea et al. 2002, Hewett & Foltz 2003, Reichard et al. 2003b), on détecte des absorptions très

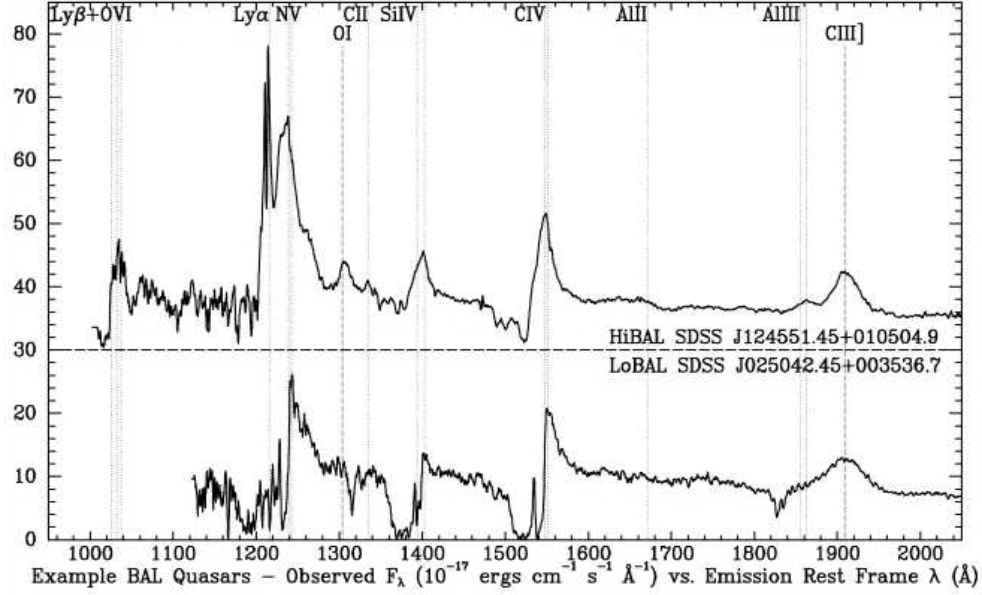


FIG. 3.6 – Exemple de deux spectres issus du SDSS (Schneider et al. 2003). La faible absorption en Lyman- $\alpha$  montre que le gaz est très ionisé dans le BAL qualifié de HiBAL, tandis que dans le LoBAL présente en plus des absorptions en basse ionisation (une forte absorption en Lyman- $\alpha$  et des absorptions en Al III et C II). Issue de Hall et al. (2002).

larges (BAL) de l'ordre de dizaines de milliers de  $km\ s^{-1}$ . Il est très probable que ces phénomènes se produisent aux environs immédiats du noyau, seul capable d'accélérer la matière aux vitesses observées. Ces absorptions sont toujours dans le bleu de l'émission du quasar, ce qui implique que le gaz responsable de ces absorptions est éjecté du centre de l'émission vers l'extérieur. Les vitesses mises en jeu ainsi que le haut degré d'ionisation indiquent que ces systèmes sont associés aux régions proches du noyau. Ces absorptions sont parfois variables comme montré par Barlow et al. (1992) par exemple. Les quasars dits 'BAL' sont classifiés en trois sous-catégories selon le gaz produisant les absorptions. La figure 3.6 montre l'exemple de deux classes de ces objets. Les BAL haute ionisation présentant des absorptions liées à des raies de haute ionisation, typiquement C IV, appelées HiBAL. Ceux de base ionisation (LoBAL) qui présentent à la fois des absorptions en haute et basse ionisation telle Mg II. Enfin les quasars de basse ionisation avec des absorptions en Fe II ou Fe III sont une sous-catégorie (FeLoBAL).

Certains systèmes d'absorption sont détectés à des décalages spectraux plus grands que  $z_{em}$ , avec des vitesses de l'ordre de  $3000\ km\ s^{-1}$ , qui sont dus à du gaz en mouvement vers le noyau.

## 3.2 Introduction à la sélection de quasars

Avant de détailler les différentes méthodes qui sont actuellement utilisées pour sélectionner les quasars, je vais présenter quelques idées simples sur leur couleur et la différenciation des objets via leurs caractéristiques photométriques. Les magnitudes d'objets obtenues dans divers filtres permettent de construire l'espace des couleurs, où chaque dimension est une différence de magnitude. Chaque objet voit sa position dans cet espace déterminé par ses caractéristiques spectroscopiques. Ainsi selon la nature de l'objet, en particulier s'il possède des raies d'émission, il aura une place bien spécifique dans un espace de couleurs donné.

Par exemple à chaque type d'étoile correspond une température particulière et des absorptions spécifiques, voire pour certaines des raies en émission. Ainsi chaque type se retrouvera dans des régions distinctes de l'espace des couleurs ; plus les différences spectrales seront grandes, plus les lieux dans l'espace des couleurs seront éloignés. Par exemple la figure 3.7 montre dans le plan  $u - g$  en fonction de  $g - r$  (couleur SDSS), des quasars à différents décalages spectraux ainsi que différents types d'étoiles.

Ce lien entre les couleurs observées et les propriétés des objets permet de séparer les étoiles et les quasars. En effet, les premiers ont typiquement un spectre de corps noir parcouru d'absorptions, tandis que le continu des quasars est une loi de puissance sur laquelle sont superposées de fortes raies d'émission. Pour utiliser cet espace des couleurs, il suffit d'y définir des régions qui doivent contenir les objets recherchés, dans notre cas les quasars. La figure 3.7 montre dans le plan  $u - g$  en fonction de  $g - r$  (couleurs SDSS) les couleurs des quasars en fonction du décalage spectral ainsi que les couleurs des étoiles. On voit bien que via les simulations réalisées grâce à des spectres composites il est possible de différencier les quasars des étoiles. Ces simulations utilisent un spectre de quasar (une DSE provenant des spectres collectés via le SDSS et le HST) à partir duquel je calcule les couleurs pour différents décalages spectraux de cette DSE en tenant compte de l'absorption par le milieu intergalactique donnée par Madau 1995. Les simulations d'étoiles sont réalisées de la même façon que dans la section 2.2.2 et je décrirai plus précisément dans la section 4.1 comment est réalisé l'ensemble de ces simulations. Néanmoins il est possible que les couleurs soient très similaires ; la superposition d'un spectre de quasar, celui d'une étoile et les filtres du SDSS (figure 3.8), permet de rendre compte que la différenciation de ces objets via leurs couleurs dans ces filtres sera très difficile.

Lors de la sélection des quasars, il convient de choisir entre sélectionner un maximum de quasars mais en ayant beaucoup de contaminants dans sa sélection et choisir peu d'objets mais avec un échantillon peu contaminé par les étoiles. Pour se faire une idée simple du fonctionnement de la sélection par diagramme couleur-couleur, il suffit de considérer une seule couleur, par exemple  $U - V$ , comme l'a fait l'équipe du sondage 'Bright Quasar Survey Sample'. En effet, la forme du continu des quasars entraîne un excès dans cette couleur. La figure 3.9 représente les coupes de sélection réalisées par l'équipe du 2dF, qui représente déjà une évolution par rapport aux simples excès  $U - V$  (utilisation de 2 couleurs au lieu d'une seule).

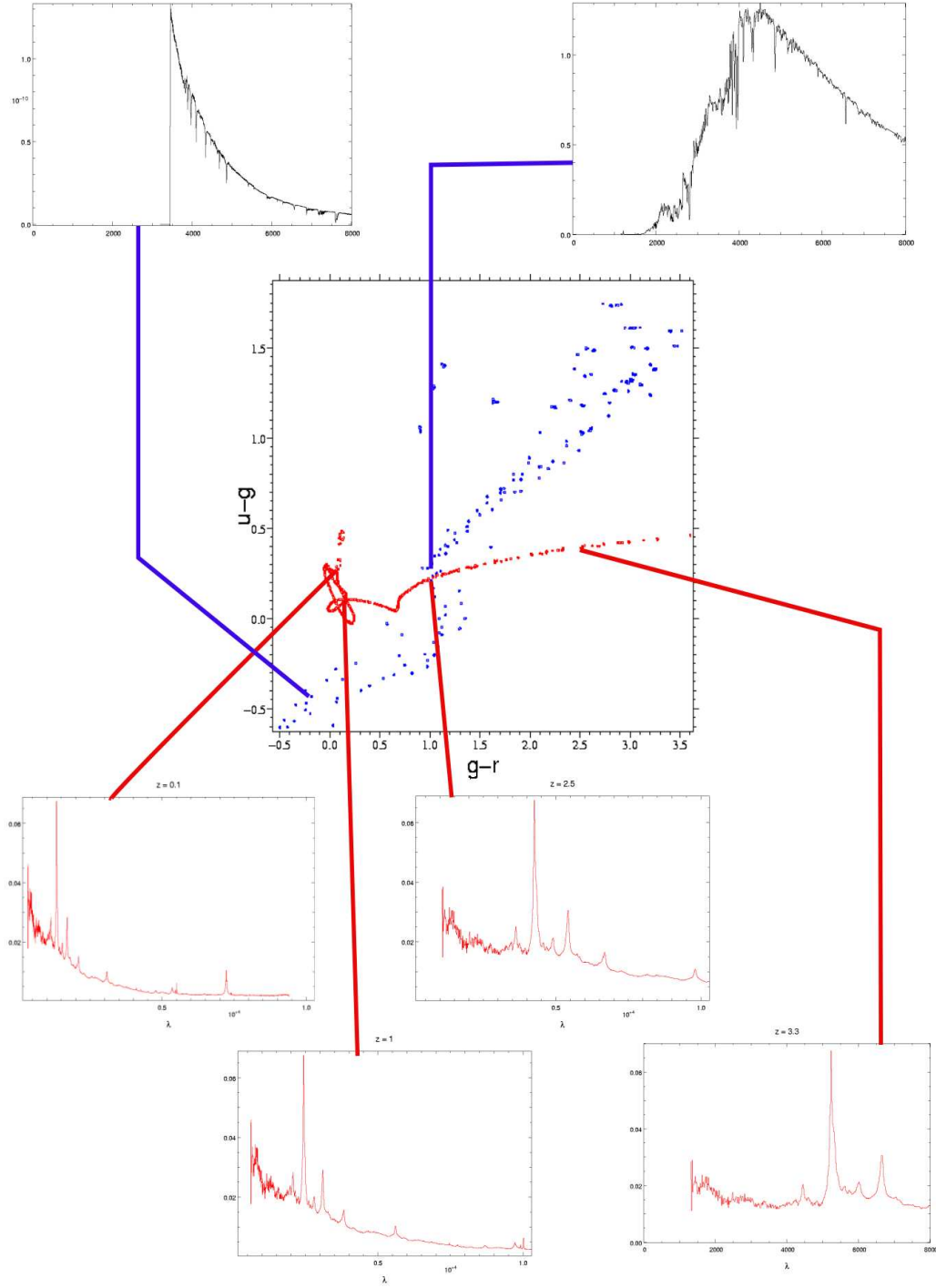


FIG. 3.7 – Au milieu, évolution des couleurs des quasars en fonction de leur décalage spectral (rouge) et répartition des étoiles dans le même plan de couleur (bleu). A titre d'exemple sont figurés les spectres de quasars (en bas, à  $z=0.1$ ; 1; 2.5 et 3.3) et d'étoiles (en haut) avec leur position dans le plan de couleur.

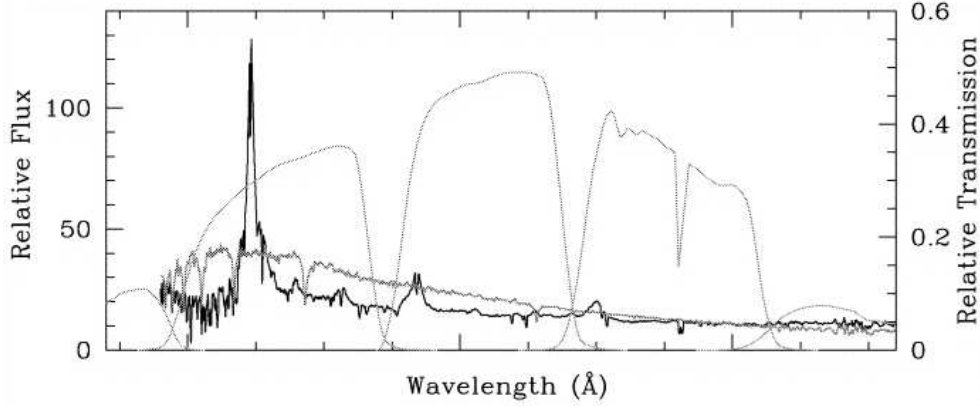


FIG. 3.8 – Spectres d'un quasar à un décalage spectral de 2.7 et une étoile avec des couleurs similaires. Le tout est superposé aux courbes de réponse des filtres du SDSS. Notez que la couleur  $g-r$  est quasiment la même pour les deux objets. Issue de Richards et al. (2002).

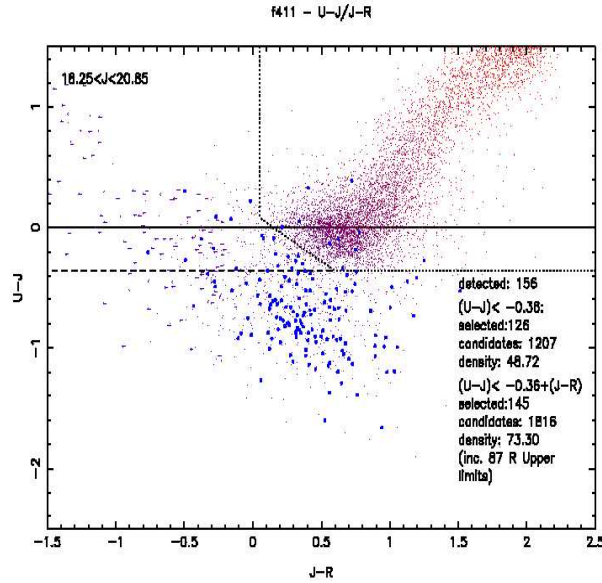


FIG. 3.9 – Évolution des critères de sélection, du simple  $U - bJ < -0.4$  aux régions définies pour le 2dF. Les quasars qui sont au-dessus du trait en pointillés n'étaient pas sélectionnés avant le 2dF. Issue du site du 2dF : <http://2dfquasar.org>.

### 3.3 Diagrammes couleur-couleur

J'ai décrit dans l'introduction les différentes méthodes qui sont disponibles actuellement pour retrouver les quasars, et ce dans différents domaines d'énergie. Ici je vais m'appliquer à décrire les méthodes de sélection dans le domaine optique. Plus particulièrement la méthode dite de diagramme couleur-couleur

qui permet d'isoler différentes régions dans l'espace des couleurs, et dans la prochaine section je décrirai la méthode d'adaptation à des Distributions Spectrales d'Energie (DSE).

Pour utiliser l'espace des couleurs, nous avons vu qu'il suffit d'y définir des régions qui doivent contenir les objets recherchés. L'idée la plus simple est de définir de façon empirique, en utilisant les données acquises par d'autres relevés spectroscopique, des régions via la projection de l'espace des couleurs dans des plans. En effet, en découpant certains plans par des droites, on définit des régions qui vont nous permettre d'exclure ou d'inclure des objets de notre sélection. En utilisant plusieurs plans des couleurs, on en vient à définir les critères de sélection. Il est aussi possible de définir des régions plus complexes, un exemple en est la définition du locus des étoiles par des ellipsoïdes et des cylindres dans l'espace des couleurs, les quasars étant alors sélectionnés comme étant à l'extérieur de ce locus. La construction de ce locus sera présentée dans la prochaine section.

Afin de distinguer des objets, il est possible de définir l'espace qu'occupent tous les objets connus et que l'on veut discerner. Dans un premier temps, divisons-les en trois grands groupes : les étoiles, les galaxies et les quasars. Comment définir les lieux dans lesquels se situent chacun de ces objets dans l'espace des couleurs ?

### 3.3.1 Le locus des étoiles dans les diagrammes couleur-couleur

Les étoiles étant relativement proches de nous, les observations de ces objets se succèdent depuis plusieurs siècles. Notre connaissance de ces objets est donc particulièrement riche, nous comprenons pour une grande partie les phénomènes qui leur permettent de rayonner, et ce jusqu'à tenter de le reproduire sur Terre. De plus chaque étoile peut être aisément identifiée et classifiée grâce à son spectre. L'étude statistique via les observations est donc aisée, de plus nous pouvons créer des modèles pour chaque type d'étoiles (Allard 1996, Pickles 1998). Ces modèles ne sont pas complètement fiables pour les étoiles froides et les très chaudes.

Ainsi la caractérisation de la position des étoiles dans les diagrammes couleur-couleur peut être conduite d'au moins deux façons différentes :

i) Tout d'abord suivant une approche empirique, (e.g. Newberg & Yanny 1997), on peut obtenir à partir d'un échantillon d'étoiles observées une paramétrisation de l'espace qu'elles occupent (à partir de données de Rich Kron, communication privée). Leur travail a été un prélude à celui de Fan 1999 (en ii) qui retrouve les mêmes résultats, je vais donc décrire le travail réalisé par ce dernier en précisant les points communs.

ii) Ou bien à travers une simulation, (e.g. Fan 1999 'Simulation of stellar objects in SDSS color space', référencé plus loin comme FanSim), qui simule les couleurs de différents objets, dont les étoiles, et les compare aux résultats de Newberg et Yanny. L'étude menée par Fan étant préliminaire à la sélection dans le SDSS (Sloan Digital Sky Survey) et s'insérant dans le cadre des tests que j'ai été amené à conduire, je vais détailler une partie de son travail.

Pour simuler les positions des différentes étoiles dans les diagrammes couleur-couleur (sans pour autant paramétrer le locus des étoiles), Fan (1999) procède

en différentes étapes :

1. Tout d'abord le nombre total d'étoiles de chaque population est calculé à partir de modèles galactiques tenant compte du disque et du halo, c'est-à-dire de leur position, pour déterminer quelle population contaminera la sélection et dans quelle proportion.
2. Ensuite pour chaque étoile il génère son âge, sa distance, sa métallicité et sa masse à partir d'un modèle de population tenant compte de la fonction de masse initiale, du taux de formation d'étoiles ainsi que de la relation entre âge et métallicité.
3. Les couleurs du SDSS et UBVRI sont calculées pour chaque étoile à partir de librairies d'étoiles. Il utilise les modèles d'atmosphère stellaire de Lejeune et al. (1997) pour calculer les magnitudes à partir de la température effective  $T_{eff}$ , de la métallicité  $[Fe/H]$  et de la gravité de surface  $\log g$ . Les modèles d'isochrones selon Bertelli et al. (1994) sont aussi utilisés pour créer des catalogues d'étoiles d'une population donnée.
4. Enfin l'extinction galactique et les erreurs photométriques sont ajoutées.

Les résultats des simulations réalisées par Fan permettent de distinguer différentes régions occupées par les étoiles, celles-ci sont représentées figure 3.10.

Dans un premier temps les simulations ont permis de voir (cf. figure 3.10) dans quelles régions se situaient les différents types d'étoiles, certaines ayant des couleurs plus ou moins proches de celles des quasars. C'est notamment le cas des étoiles bleues, dont certaines sont présentes dans le halo. Dans le plan  $g - r$  en fonction de  $u - g$ , les étoiles du halo pauvres en métaux, représentées par des cercles, apparaissent dans la partie bleue du plan (c'est-à-dire en bas à gauche), en opposition aux étoiles du disque symbolisées par des croix. Le plan  $g - r$  en fonction de  $r - i$  permet de distinguer en plus de ces étoiles les étoiles carbonées (issues d'observations par Krisciunas et al. 1998) et celles qui sont entourées de poussières qui possèdent une couleur  $g - r$  plus rouge, due à un rougissement interne différent des étoiles plus froides. Toutes ces remarques, si elles nous donnent une idée de façon assez intuitive des lieux où se trouvent les étoiles et quels sont les éventuels contaminants aux sélections de quasars, n'apportent toutefois pas de définition claire des régions qu'occupent les étoiles. Il faut donc, à partir des simulations et des observations, définir de façon systématique le locus des étoiles, c'est-à-dire lui attribuer, dans l'espace des couleurs, des ellipsoïdes clairement définis. Afin de sélectionner les quasars, qui se situeraient à l'extérieur du locus des étoiles, l'équipe du SDSS a choisi dans un premier temps d'utiliser une définition du locus des étoiles à partir des premières données du SDSS (Richards et al 2002). Pour ce faire, ils ont appliqué les mêmes techniques que Newberg & Yanny (1997). La construction du locus a consisté en la définition des paramètres d'ellipses cylindriques qui contiennent les régions les plus denses d'étoiles. Les étapes de la sélection du SDSS seront décrites dans la partie 3.3.4, c'est aussi dans cette partie qu'est présenté plus en détail le locus des étoiles, représenté figure 3.15.

Les étoiles occupent une région quasiment à une dimension dans l'espace  $(u - g)$ ,  $(g - r)$ ,  $(r - i)$ ,  $(i - z)$  avec comme paramètre principal de position le long de ce locus, la température. Bien sûr ce locus ne contient pas toutes les étoiles, mais les



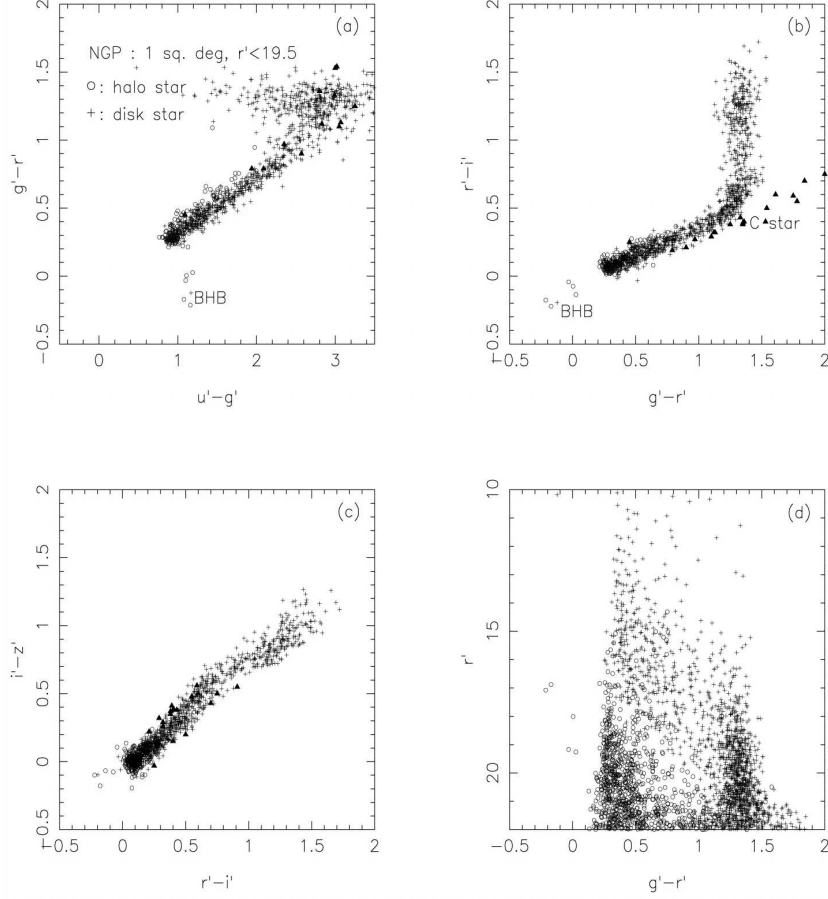


FIG. 3.10 – Diagramme couleur-couleur et couleur-magnitude à partir d'une simulation d'un degré carré du pôle galactique nord. Les triangles sont des observations d'étoiles carbonées par Krisciunas et al. (1998). Issue de Fan (1999).

plus ordinaires, de type F ou M qui dominent. La recherche d'objets à l'extérieur de ce locus a permis de déceler des objets aux caractéristiques particulières (e.g. Bazell et al. 2005). Donc la plupart des étoiles peuvent se retrouver dans un locus qui est à définir (cf. la partie 3.3.4), mais il existe d'autres contaminants qui doivent être traités différemment, notamment les naines blanches, cf. section 3.3.5.

### 3.3.2 Position des galaxies dans les diagrammes couleur-couleur

Si l'on restreint la sélection aux objets ponctuels, on ne devrait pas se préoccuper a priori des couleurs des galaxies. En effet, le nombre de galaxies est faible parmi les objets ponctuels du relevé mais néanmoins certaines sont présentes. Ces galaxies sont qualifiées de galaxies compactes à raies d'émission (CELG). Ces objets ont des raies d'émission très étroites et un continu en loi de puissance (dû à une activité forte de formation d'étoiles), ce qui leur confèrent des couleurs très similaires à celles des quasars à bas décalage spectral. Fan trouve que

les couleurs de ces galaxies dans des filtres larges tels que ceux du SDSS sont essentiellement dus à leur continu, les raies peu larges jouant un rôle mineur. Ce continu est plus fort dans le rouge que celui des quasars, ainsi la couleur est plus rouge, ce qui est la clé pour pouvoir les distinguer. Néanmoins une fraction de ces galaxies ont un continu relativement faible et des raies en émission très fortes, ce qui implique de fortes variations de couleur comparées à la loi de puissance et qui les situent au final dans des régions très éloignées des étoiles. Il faut donc s'attendre malgré une sélection sur des objets ponctuels à retrouver parmi les contaminants des galaxies à raies d'émission. Sur les tests que j'ai menés (cf. section 4.2), les galaxies restent une minorité des contaminants  $\sim 5\%$ . A l'inverse, pour donner quelques chiffres sur les quasars résolus, à partir du catalogue de Schneider et al. (2005), 2077 sont classifiés comme étant des objets non ponctuels (soit 4%). La plupart d'entre eux ont un décalage spectral inférieur à 1 (94%), 7 sont au-dessus de 3. La majorité de ces quasars à grand décalage spectral correspondent probablement à des erreurs de mesure. Toutefois certains d'entre eux peuvent être des superpositions fortuites d'objets en avant plan ou de possible images, issues de lentille gravitationnelle, très faiblement séparées (Schneider et al. 2005).

### 3.3.3 Position des quasars dans les diagrammes couleur-couleur

L'étude menée par Fan (1999) avait pour but de déterminer les régions de sélection des quasars dans le 'early data release' du SDSS (Stoughton et al. 2002). La détermination de ces régions a nécessité de simuler les couleurs des quasars mais aussi des étoiles et des galaxies compactes à raie d'émission. Lors de la simulation des couleurs des quasars ; Fan a pris en compte les trois points suivants dans ses simulations pour modéliser les quasars :

1. Le continu en loi de puissance en optique et ultra-violet. La distribution des indices spectraux des lois de puissance pour le continu d'un quasar a été modélisée par une gaussienne centrée sur l'indice de loi de puissance (issus de Francis et al. 1991) de  $0.5 \pm 0.3$ .
2. En ce qui concerne les raies en émission : les rapports entre les seize principales raies d'émission sont issues des catalogues de Wilkes (1986) et de Francis et al. (1991). La valeur précise des largeurs des raies n'a que peu d'influence sur les couleurs en photométrie issue de filtres à large bande.
3. Prenant en compte aussi des systèmes d'absorption pour les quasars à grand décalage spectral suivant Fan & Chen (1993) et Warren et al. (1994).

En ce qui concerne le dernier point, si les étoiles voient leur position évoluer dans les diagrammes couleurs à cause de leur type, c'est-à-dire entre autre aux phénomènes qui sont à la source de leur rayonnement, pour les quasars c'est leur éloignement par rapport à l'observateur qui joue un rôle essentiel. En effet, les raies d'émission et le continu sont décalés dans le rouge au fur et à mesure que le quasar est éloigné de nous, ce qui influe sur les couleurs observées. La figure 3.12 montre l'évolution des couleurs supposées des quasars en fonction du décalage spectral dans différents plans de couleurs.

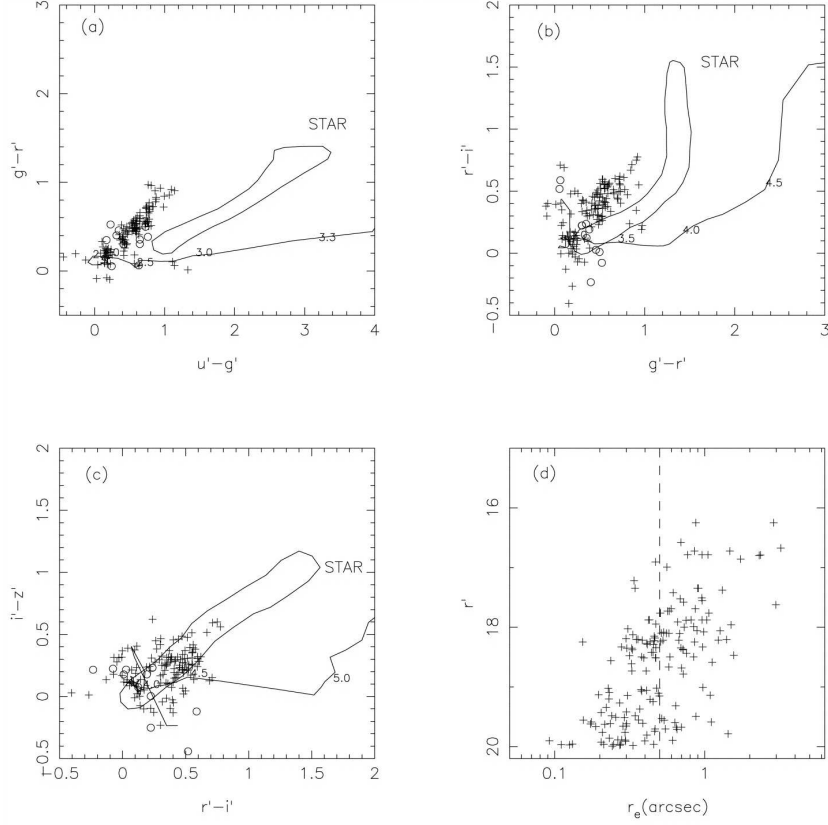


FIG. 3.11 – Diagrammes couleur-couleur et distribution du rayon des galaxies compactes simulées, représentées par les signes plus. Les couleurs des galaxies venant des observations sont issues de Hall et al. (1996), représentées par des cercles. Sont représentés le locus des étoiles et le lieu de quasars à différents décalages spectraux. Issue de Fan (1999).

La position des quasars via les simulations est représentée figure 3.13. Ceci permet de se rendre compte de l'évolution des couleurs en fonction du décalage spectral. Il est clairement visible que pour des décalage spectraux supérieurs à 3 la différenciation étoiles-quasars sera plus facile. Les figures 3.12 et 3.13 rendent compte de plusieurs points essentiels.

Pour des quasars situés à un décalage spectral inférieur à 2, la couleur  $u - g$  reste inférieure à 0.3, ceci étant dû au fait qu'elle témoigne essentiellement (pour cet intervalle en décalage spectral) du continu et pas des raies d'émission, leur largeur équivalente étant trop faible pour induire de fortes variations de magnitude. Ceci s'observe jusqu'à un décalage spectral de 2, c'est-à-dire jusqu'à ce que l'émission Lyman- $\alpha$  rentre dans le domaine couvert par le filtre  $u$ , ce qui a pour effet de rendre la couleur  $u - g$  bien plus bleue. Cette couleur bleue peut aussi caractériser certaines galaxies compactes à raie d'émission qui viennent contaminer l'échantillon de candidats.

Au-delà d'un décalage de 2.5, les systèmes d'absorptions présents dans la région bleue de l'émission Lyman- $\alpha$ , la forêt Lyman- $\alpha$ , rentrent dans le domaine

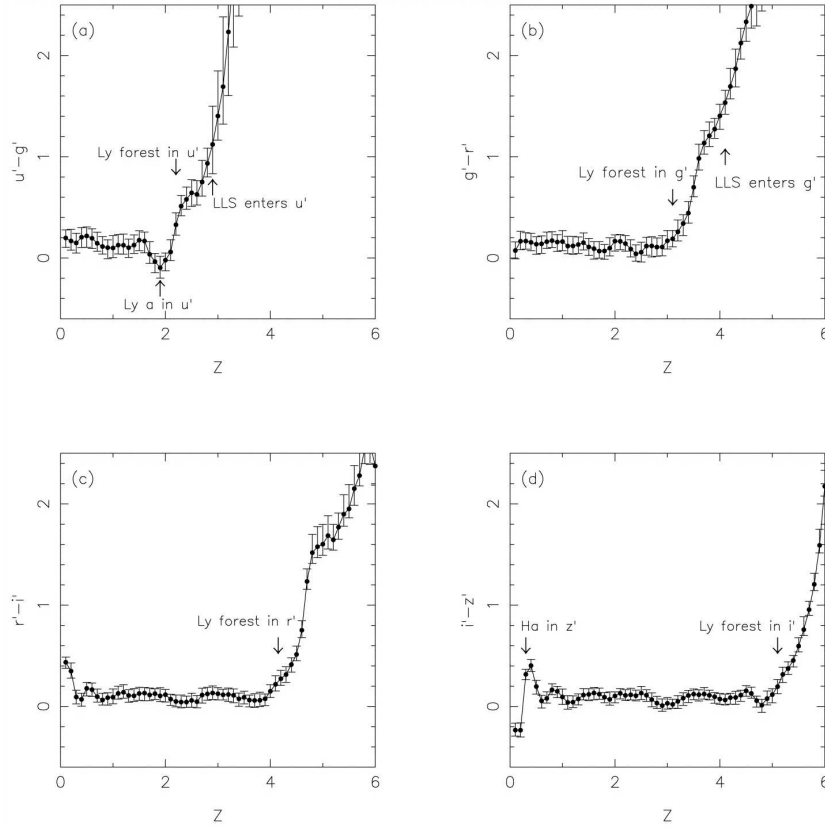


FIG. 3.12 – Évolution des couleurs des quasars en fonction du décalage spectral. La figure montre la médiane obtenue pour une simulation de 100 quasars à chaque décalage spectral. Issue de Fan (1999).

du filtre  $u$ , puis les systèmes à la limite de Lyman (LLS, pour Lyman limit system, qui sont "des systèmes qui possèdent une densité de colonne en hydrogène neutre suffisante pour qu'une discontinuité soit observée dans le continu du quasar à  $(1 + z_{abs}).912\text{\AA}$ ", Petitjean 1999) ce qui a pour effet de rougir la couleur  $u - g$ .

Le comportement des autres couleurs est sensiblement le même, mais pour les plus grands décalages spectraux. Les quasars à un décalage spectral entre 2.5 et 3 ont des couleurs qui sont très proches de celles des naines blanches, des étoiles bleues du halo et de quelques galaxies à raies d'émission. Ceci se voit dans une région de l'espace des couleurs où  $u - g \sim 0.8$ ,  $g - r \sim 0.3$ , et  $r - i$ ,  $i - z \sim 0$ . Cet intervalle est celui qui contient le plus de contaminants pour la sélection de quasars. Mais c'est aussi celui qui présente le plus d'intérêt pour la tomographie du milieu intergalactique, puisque celle-ci nécessite la présence de la forêt Lyman- $\alpha$  dans les spectres des quasars qui seront pris depuis des observatoires terrestres. Au-delà d'un décalage spectral de 3, les quasars ont des couleurs particulièrement rouges qui les rend faciles à distinguer (mais ils sont bien moins nombreux).

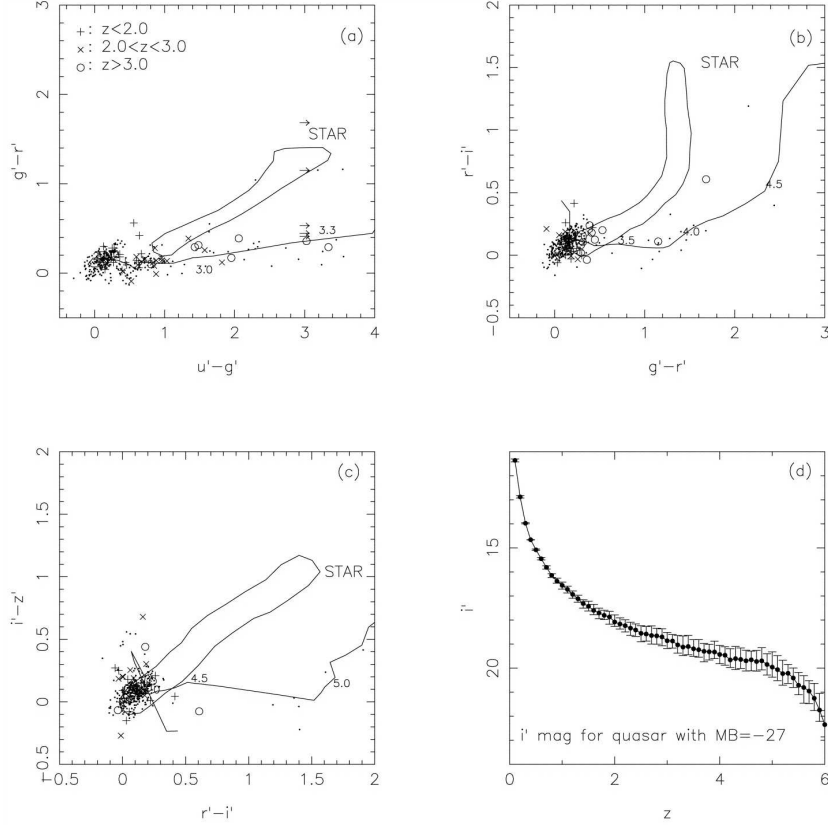


FIG. 3.13 – Panneau (a), (b) et (c) Diagrammes couleur-couleur des quasars à différent décalages spectraux simulés avec  $i < 19.5$  (point noirs) ainsi que des observations issues du SDSS (symbole plus grand). Les flèches symbolisent les limites inférieures en  $u - g$  dues aux non-détections dans le filtre  $u$ . Les lignes représentent le chemin suivi par les couleurs des quasars en fonction de leur évolution en décalage spectral, comme précisé dans la figure 3.12. Le locus des étoiles est aussi représenté. Panneau (d) : médiane des quasars simulés avec une magnitude intrinsèque similaire à celle de 3C 273 en fonction du décalage spectral. Issue de Fan (1999).

Une fois le 'early data release' du SDSS complété, l'équipe de SDSS a décidé de raffiner sa sélection, car l'échantillon de candidats contenait trop de contaminants, ceci étant notamment dû à un problème sur la localisation des quasars dans les diagrammes couleur-couleur.

Le second algorithme ainsi développé pour la sélection de quasars dans le SDSS a consisté à mettre en place le locus des étoiles, et à extraire les objets qui n'appartenaient pas à ce locus pour ensuite appliquer des coupes dans l'espace des couleurs. La dernière sélection du SDSS néanmoins met à profit le nombre de quasars déjà trouvés par les premières observations du SDSS. Ainsi, de la même façon qu'ils ont défini le locus des étoiles, Richards et al. (2004) définissent le locus des quasars. Ces régions définies doivent contenir un maximum de quasars avec un minimum de contaminants. Ces sélections sont détaillées dans la section

suivante (3.3.4).

### 3.3.4 Algorithmes de sélection des quasars dans le SDSS

Les sélections par diagramme couleur-couleur ont atteint leur dernier développement avec les algorithmes du SDSS. Je vais présenter deux étapes clés de leur évolution, tout d'abord la première sélection publiée par Richards et al. (2002) et ensuite le dernier catalogue de candidats grâce à une évolution de l'algorithme publié par Richards et al. (2004).

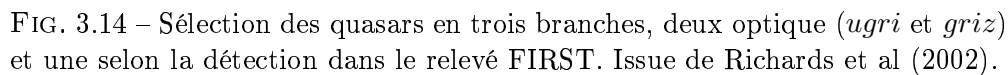
#### Premier algorithme de sélection du SDSS

La sélection des candidats pour les cibles spectroscopiques du relevé SDSS (cf. section 2.1) peut se décomposer en deux parties, l'une utilisant les couleurs basées sur les filtres *ugri* et l'autre sur les filtres *griz*, les détails pouvant être trouvés dans Richards et al. (2002). La sélection dans le SDSS est essentiellement basée sur deux branches de sélection tridimensionnelle (3D) dans l'espace des couleurs, il s'agit bien d'une sélection directe dans l'espace, sans projection donc. La première branche utilise les filtres *ugri* pour identifier les quasars avec un excès *UV* tandis que la seconde utilise les filtres *griz* pour identifier les quasars de décalage spectral supérieur à 3 (cf. figure 3.14). À noter que lors de cette sélection, il n'a pas été fait de différence entre les quasars et les NAG, les galaxies de Seyfert par exemple étant considérées comme étant des quasars à faible décalage spectral. La sélection a donc porté sur les objets ponctuels et étendus.

Le code de sélection fonctionne comme suit :

1. Les objets qui sont connus comme ayant des problèmes de flux, dans leur calibration ou au niveau de leur acquisition dans la partie imagerie sont rejetés (cf. 2.1).
2. Les sources ponctuelles qui correspondent à des sources FIRST, en radio donc, sont sélectionnées sans passer par leurs couleurs, je ne m'attarderai donc pas dessus.
3. Les autres sources sont comparées à une distribution d'étoiles et de galaxies dans deux espaces des couleurs distincts, un pour les quasars à bas décalage spectral (*ugri*) et un autre pour les quasars à plus grand décalage (couleurs *griz*). Simplement, les étoiles suivent un locus particulier et les objets en dehors de celui-ci sont sélectionnés pour un suivi spectroscopique, après avoir subi un découpage dans l'espace des couleurs. C'est cette partie que je vais détailler.

Il est donc important de comprendre que la première difficulté de la sélection est la construction du locus des étoiles, car ce sont les objets en dehors de ce locus (en dehors signifiant à une certaine distance des limites de ce lieu, tenant compte des erreurs associées à chaque objet) qui feront l'objet des découpages



Les magnitudes limites et donc le nombre de candidats qui seront pris en compte dépendent en grande partie des capacités du suivi spectroscopique. En effet, le spectrographe du SDSS peut dédier 18 de ses fibres par degré carré au suivi spectroscopique des candidats quasars. En fonction de cette contrainte, de l'efficacité et de la complétude attendue, la sélection s'est faite sur les candidats

quasars ayant une magnitude en  $i$  corrigée de l'extinction galactique supérieure à 15 et inférieure à 19.1 pour la sélection à bas décalage spectral et à 20.2 pour les quasars à grand décalage spectral (supérieur à 3 typiquement).

Il est important de noter que cet algorithme a été construit après que des suivis spectroscopiques aient été conduits, ainsi leur travail a pu être enrichi de données spectroscopique. La base de données spectrale de quasars a été construite pour arriver aux buts d'efficacité et de complétude qu'ils s'étaient fixés (Richards et al. 2002 section 2).

La création du locus des étoiles s'est faite quasiment de la même façon que pour Newberg & Yanny (1997), Fan (1999) et Finlator et al. (2000). La différence réside dans le fait que ces derniers travaillent dans l'espace des couleurs comptant quatre dimensions tandis que dans le cas de cette sélection le code a fonctionné sur deux espaces différents à trois dimensions ceci afin de simplifier l'algorithme. Les détails de la création du locus peuvent être trouvés dans Richards et al. (2002).

Les objets sont sélectionnés en 4 parties, i) une concernant l'espace des couleurs *ugri* exclusivement cherchant les quasars à bas décalage spectral  $i < 19.1$ , ii) une autre dans l'espace *griz* cherchant les quasars à grand décalage spectral jusqu'à  $i < 20.2$ , iii) des zones d'exclusion d'objets, iv) des zones d'inclusions.

Les objets qui vont faire l'objet des sélections i) et ii) ont d'abord été sélectionnés comme étant à l'extérieur du locus des étoiles dans chacun des espaces de couleurs respectifs. D'une façon simple, ceci peut être dit comme suit : "We select all objects whose photometric error ellipsoids (defined by  $N_{\sigma_{err}}$  times the one-sigma error ellipsoid) do not intersect the stellar locus." (Richards et al. 2002, Appendice B). La construction des ellipses d'erreur pour chaque objet suppose que les erreurs dans chaque filtre sont non-corrélées, ce qui n'est pas tout à fait exact, cf. Scranton et al. (2005). En résumé, **tous les objets sélectionnés sont d'abord choisis à l'extérieur du locus des étoiles**. Les projections de ce locus ainsi que sa paramétrisation sont représentées figure 3.15. Il est important de noter que tous les objets du 'early data release' ont été utilisés pour réaliser ce locus. C'est-à-dire que tous les objets ponctuels ont été pris pour cette paramétrisation, même ceux sans confirmation spectroscopique. Ceci est possible car dans les régions qu'ils utilisent, la majorité des objets sont des étoiles. L'équipe du SDSS exclut néanmoins de ce locus les régions de l'espace telles que :  $u - g < 0.5$  ou  $g - r < 0$  qui regroupent essentiellement des quasars et des étoiles de type A. Les ellipses en rouge montrent la projection de la surface à 4-sigma. En pratique des objets avec une erreur nulle seraient sélectionnés à partir de cette surface. En réalité, en tenant compte des erreurs sur les magnitudes, les régions d'exclusion sont plus grandes.

#### i) Sélection dans l'espace des couleurs *ugri* :

La sélection des quasars dans l'espace des couleurs *ugri* utilise une distinction de traitement entre objets ponctuels ou étendus. Les objets étendus détectés en  $u$  et  $g$ , ayant des erreurs inférieures à 0.2 et dont la couleur  $u - g$  est strictement supérieure à  $> 0.9$  sont rejetés, ainsi que ceux ayant  $l > 0$  et  $k > 0$  selon les notations de Newber & Yanny (1997) (voir aussi Richards et al. (2002) annexe



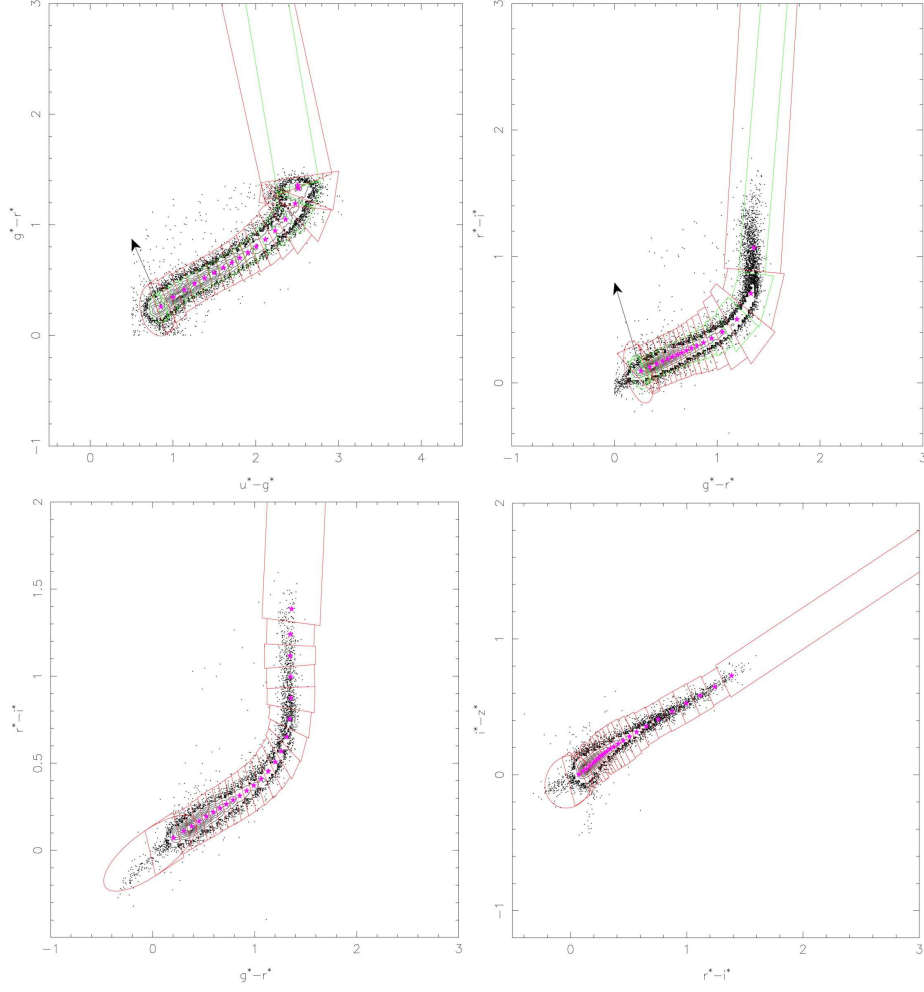


FIG. 3.15 – Diagrammes couleur-couleur représentant la projection du locus des étoiles dans différents plans de couleurs. Les objets utilisés pour sa caractérisation sont montrés en points et contours noir, soit près de 50 000 points. Les symboles en étoile magenta représentent la trace du centre du locus. Les lignes rouges représentent les limites d’une surface à 4-sigma autour de ces centres. Les régions en vert sont les limites de sélection des objets extérieur au locus (qui n’est plus à 4-sigma) pour les quasars avec des décalages spectraux compris entre 2 et 3 (cf. paragraphe sur les régions d’inclusion). Issue de Richards et al. (2002).

A, la signification de ces paramètres n’est pas l’objet de notre développement). Les objets qui répondent à ces conditions et ne se situent pas dans les régions d’exclusion qui seront données dans un prochain paragraphe sont sélectionnés comme candidats quasars.

## ii) Sélection dans l’espace des couleurs *griz* :

Ces critères visent les quasars à plus grand décalage spectral ; pour éviter de

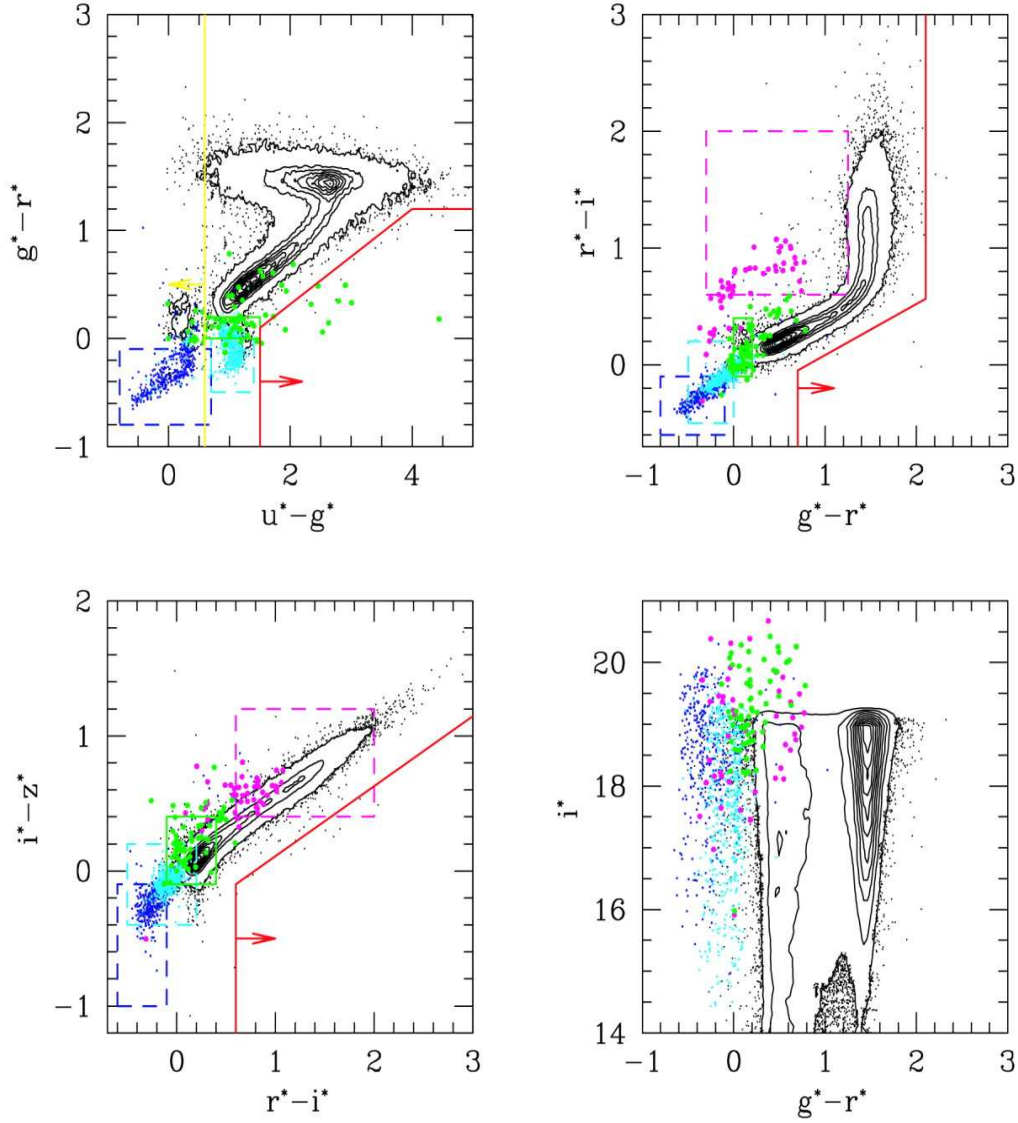


FIG. 3.16 – Régions d'exclusion et d'inclusion pour la sélection de quasars. Les points noirs et les contours sont les sources ponctuelles avec  $i < 19.1$ . Les lignes en bleu sombre et tirets montrent les régions d'exclusion des naines blanches, les points en bleu sombre sont les naines blanches confirmées par spectroscopie. De même pour les étoiles de type A en bleu clair et tiret, ainsi que pour les paires naine blanche - étoile de type M en magenta et tirets et aussi pour les quasars aux décalages spectraux entre  $2.5 < z < 3.0$  en vert et ligne pleines. Notez que la boîte verte n'est pas une boîte d'exclusion mais d'inclusion et que les points verts représentent **tous** les quasars avec  $2.5 < z < 3.0$ , pas seulement ceux qui sont dans la boîte. De plus ces boîtes sont les projections des boîtes réelles dans l'espace à quatre dimensions. La ligne jaune montre la coupe en couleur UVX. Les lignes rouges montrent les différentes coupes pour les grands décalages spectraux. Issue de Richards et al. (2002).

sélectionner des quasars à bas décalage spectral dans cette sélection, les objets répondant aux critères suivants ne sont pas choisis comme quasars :

$$\begin{aligned} \text{A)} \quad & g - r < 1.0 \\ \text{B)} \quad & u - g \geq 0.8 \\ \text{C)} \quad & i \geq 19.1 \quad \text{OR} \quad u - g < 2.5. \end{aligned} \tag{3.2}$$

### Régions d'exclusion :

Il y a trois régions d'exclusion suivant les types d'objets à exclure : les naines blanches, les paires naine blanche - étoile de type M et les étoiles de type A. La région d'exclusion des naines blanches, est figurée en bleu foncé dans le diagramme 3.16 et répond à l'intersection des critères ci-dessous :

$$\begin{aligned} \text{A)} \quad & -0.8 < u - g < 0.7 \\ \text{B)} \quad & -0.8 < g - r < -0.1 \\ \text{C)} \quad & -0.6 < r - i < -0.1 \\ \text{D)} \quad & -1.0 < i - z < -0.1. \end{aligned} \tag{3.3}$$

Ces limites laissent la possibilité de sélectionner des objets ayant une couleur bleue particulièrement importante, sans en connaître la nature ; ce qui laisse la place pour la découverte d'objets inhabituels.

Ensuite les régions couvrant les étoiles de type A, trop peu nombreuses pour être prises en compte lors de la définition du locus, mais qui peuvent constituer une part significative des contaminants, ces régions sont figurées en bleu clair dans le diagramme 3.16 et sont définies comme étant l'intersection des régions :

$$\begin{aligned} \text{A)} \quad & 0.7 < u - g < 1.4 \\ \text{B)} \quad & -0.5 < g - r < 0.0 \\ \text{C)} \quad & -0.5 < r - i < 0.2 \\ \text{D)} \quad & -0.4 < i - z < 0.2. \end{aligned} \tag{3.4}$$

Les dernières régions de l'espace des couleurs rejettent les objets tels que les paires d'étoiles bleues et rouges qui ne sont pas résolues et les paires d'étoiles de type M et naine blanche. Cette région est figurée en magenta dans le diagramme 3.16 et est définie comme étant l'intersection de :

$$\begin{aligned} \text{A)} \quad & -0.3 < g - r < 1.25 \\ \text{B)} \quad & 0.6 < r - i < 2.0 \\ \text{C)} \quad & 0.4 < i - z < 1.2 \\ \text{D)} \quad & \sigma_g < 0.2, \end{aligned} \tag{3.5}$$

A part la partie portant sur les erreurs en  $g$ , il n'y a pas de contraintes sur les erreurs et les régions ne concernent que les couleurs naturelles des objets.

### Régions d'inclusion :

En plus des objets qui sont rejetés, l'équipe du SDSS a défini des régions où les objets qui y sont présents sont immédiatement inclus dans la sélection même s'ils ne répondent pas aux critères définis par le locus des étoiles. Ceci est fait essentiellement afin de sélectionner les quasars à un décalage spectral de  $2.5 < z < 3$  dont les couleurs croisent le locus des étoiles. Certaines régions sont aussi définies afin de sélectionner des quasars à grand décalage spectral qui pourraient être manqués par les autres sélections.

Tout d'abord une première région de l'espace des couleurs est définie comme suit pour sélectionner les quasars entre 2.5 et 3 :

$$\begin{aligned}
 \text{A)} \quad & 0.6 < u - g < 1.5 \\
 \text{B)} \quad & 0.0 < g - r < 0.2 \\
 \text{C)} \quad & -0.1 < r - i < 0.4 \\
 \text{D)} \quad & -0.1 < i - z < 0.4.
 \end{aligned} \tag{3.6}$$

Une fois les objets sélectionnés dans cette région, ils sont restreints à être des sources ponctuelles qui sont à plus de  $2\sigma$  du locus des étoiles (représenté en vert sur la figure 3.15). Comme cette région contient aussi beaucoup d'étoiles et afin de ne pas trop diminuer l'efficacité de leur algorithme, l'équipe du SDSS a choisi de n'observer que 10% des objets récupérés via cette inclusion.

Une autre région concerne les quasars avec un décalage spectral  $z \leq 2.2$ . Les objets qui ont des erreurs en  $u$  et  $g$  inférieures à 0.1 et avec  $u - g < 0.6$  sont sélectionnés s'ils ne sont pas dans la région d'exclusion des naines blanches.

Les régions suivantes ont pour but de récupérer un maximum de quasars à grand décalage spectral. Régions définies par Fan et al. 2001 dans l'espace  $gri$  pour récupérer les quasars avec des décalages spectraux  $z \geq 3.6$ , elles correspondent aux intersections suivantes :

$$\begin{aligned}
 \text{A)} \quad & \sigma_i < 0.2 \\
 \text{B)} \quad & u - g > 1.5 \quad \text{OU} \quad u > 20.6 \\
 \text{C)} \quad & g - r > 0.7 \\
 \text{D)} \quad & g - r > 2.1 \quad \text{OU} \quad r - i < 0.44 * (g - r) - 0.358 \\
 \text{E)} \quad & i - z < 0.25 \\
 \text{F)} \quad & i - z > -1.0.
 \end{aligned} \tag{3.7}$$

Les coupes C) et D) sont représentées dans la figure 3.16, en rouge. La contrainte sur les erreurs en  $i$  A), assure l'intégrité des couleurs  $r - i$  et  $i - z$ . Les coupes en  $u - g$  et  $u$ , B), restreignent l'échantillon aux 'dropouts' dans le filtre  $u$ . La coupe verticale en  $g - r$  et celles diagonales C) et D), préviennent la région de trop se rapprocher du locus des étoiles. Enfin la restriction E) empêche la plupart des étoiles de type M d'être sélectionnées et la F) prévient les objets trop bleu en  $i - z$  d'être sélectionnés.

Les régions dans l'espace  $riz$  sont prévues pour sélectionner les quasars avec

un décalage spectral  $z \geq 4.5$  :

$$\begin{aligned}
 \text{A)} \quad & \sigma_i < 0.2 \\
 \text{B)} \quad & u > 21.5 \\
 \text{C)} \quad & g > 21.0 \\
 \text{D)} \quad & r - i > 0.6 \\
 \text{E)} \quad & i - z > -1.0 \\
 \text{F)} \quad & i - z < 0.52 * (r - i) - 0.412.
 \end{aligned} \tag{3.8}$$

Dans ce cas les limites sur  $u$  et  $g$ , B) et C), sont faites pour restreindre l'échantillon d'objets aux 'dropouts' en  $u$  et en  $g$ . La coupe en  $r - i$  D) met une limite sur la quantité de flux dans le filtre  $r$  par rapport à  $i$  et évite aux objets sélectionnés d'être trop près du locus des étoiles. La limite en bleu sur  $i - z$  E) a le même rôle que précédemment. Finalement la coupe diagonale en F) limite le nombre d'étoiles 'normales'.

La dernière inclusion concernant les couleurs  $ugr$  a pour but de récupérer les quasars avec  $z \geq 3.0$ , notamment ceux qui sont plus faibles que  $i = 19.1$ , et ceux avec  $z \sim 3.5$  qui peuvent être détectés en  $u$  :

$$\begin{aligned}
 \text{A)} \quad & u > 20.6 \\
 \text{B)} \quad & u - g > 1.5 \\
 \text{C)} \quad & g - r < 1.2 \\
 \text{D)} \quad & r - i < 0.3 \\
 \text{E)} \quad & i - z > -1.0 \\
 \text{F)} \quad & g - r < 0.44 * (u - g) - 0.56.
 \end{aligned} \tag{3.9}$$

Les coupes correspondantes sont représentées sur la figure 3.16 dans le panneau du haut à gauche. Les limites A) et B) ont pour but de s'assurer que les objets sont suffisamment rouges pour être des quasars à grand décalage spectral. La coupe F) et la limite supérieure C) gardent les objets loin du locus des étoiles. La coupe  $r - i$  D) limite le nombre d'étoile de type M. Enfin comme pour les régions précédentes, les objets trop bleus en  $i - z$  E) sont rejetés. En pratique ces coupes dans l'espace  $ugr$  sont combinées avec une autre sélection qui consiste à prendre les objets en dehors du locus des étoiles dans l'espace  $ugri$ ; ces objets doivent aussi être détectés dans  $u$  et  $g$  et avoir des erreurs inférieures à 0.2 magnitude et une couleur  $u - g > 1.5$ .

Les objets qui remplissent ces critères ou ceux décrits au-dessus sont alors sélectionnés comme étant des candidats quasars à grand décalage spectral, jusqu'à la limite de 20.2, pour peu qu'ils ne soient pas étendus.

## Seconde sélection du SDSS

La première sélection était essentiellement basée sur la caractérisation de l'espace occupé par les étoiles et par la sélection des objets se trouvant à l'extérieur de ce locus. Une autre approche possible est de déterminer quel est le locus occupé par les quasars et, une fois celui-ci déterminé, sélectionner les objets qui se trouvent à l'intérieur. Ceci ne peut être fait qu'à la seule condition que l'échantillon de quasars utilisé pour définir ce locus soit suffisamment important et a priori non biaisé. C'est ce qu'ont fait Richards et al. (2004). Je ne

vais pas donner tous les détails de l'algorithme, qui peuvent être trouvés dans les références, je vais simplement en donner quelques rapides caractéristiques.

Pour classer les objets comme étoiles ou quasars, l'équipe a utilisé deux échantillons, l'un de test et l'autre d'étalonnage. L'échantillon d'étalonnage est constitué d'étoiles et de quasars. Pour ceux-ci le catalogue de Schneider et al. (2003), constitué de 16 713 quasars, a été utilisé. Pour chaque objet qui appartient à l'échantillon de test et qui doit être identifié (en étoile ou en quasar), est associée une probabilité d'être une étoile ou un quasar, la meilleure probabilité permettant l'identification de l'objet. La probabilité d'être une étoile pour un objet  $x$  (représenté par quatre couleurs) peut être définie comme étant la vraisemblance ("likelihood function") de  $x$  comparée à une fonction de densité de probabilité qui décrit les étoiles,  $p(x|C_1)$  où  $C_1$  est la classe 1, celle des étoiles ( $C_1$  est la variable pour  $x$  fixé). Cette fonction de probabilité est déterminée par l'estimation d'une densité de noyau (Kernel Density Estimate, KDE; Silverman 1986). Pour donner une idée de sa représentation, il est possible de la figurer comme un histogramme, où, à la place de pas discrets définis par une grille, le KDE définit chaque 'pas' avec son centre et la largeur du 'pas' par une fonction continue (i.e. une fonction gaussienne en quatre dimensions). Lorsque les différents échantillons sont bien définis, il est possible de commencer le processus de classification en calculant la ressemblance de chaque objet  $x$  dans l'échantillon de test en regard de chaque échantillon d'étalonnage (i.e. la densité de  $x$  par rapport à celle des étoiles et des quasars) en utilisant le KDE (ce qui revient grossièrement à calculer une distance à tous les points de l'échantillon d'étalonnage) :

$$\hat{p}(x) = \frac{1}{N} \sum_i^N K_h(||x - x_i||) \quad (3.10)$$

Est la probabilité pour  $x$  d'être une étoile. Où  $N$  est le nombre de données,  $K_h(z)$  est la fonction du noyau qui satisfait  $\int_{-\infty}^{\infty} K_h(z)dz = 1$ ,  $h$  est un facteur de normalisation et  $z$  est la distance entre un point dans l'échantillon de test et un point dans l'échantillon d'étalonnage (dans ce cas il s'agit de distance euclidienne dans un espace des couleurs à quatre dimensions,  $||x - x_i||$ ). Le noyau utilisé est gaussien avec comme facteur de normalisation la variance de la gaussienne, i.e.  $K_h(||x - x_i||) = \frac{1}{h\sqrt{2\pi}} \exp -\frac{||x-x_i||^2}{2h^2}$ . Pour faire une analogie avec un histogramme à une dimension, celui-ci aurait un pas de largeur  $h$ , aurait comme KDE  $K_h(z) =$  une constante et le lieu des pas est fixé plutôt que centré sur chaque point des données.

Ensuite il convient d'utiliser la probabilité a priori que l'objet est une étoile, notée  $P(C_1)$ , suivant les règles de Bayes (Bayes 1763; Press et al. 1992) qui permet de pondérer les vraisemblances :

$$P(C_1|x) = \frac{p(x|C_1)P(C_1)}{p(x|C_1)P(C_1) + p(x|C_2)P(C_2)} \quad (3.11)$$

En particulier les objets avec  $P(C_1|x) > 0.5$  sont classifiés comme étant des étoiles et les objets avec  $P(C_1|x) < 0.5$  sont classifiés comme étant des quasars.  $P(C_1)$  est estimée à partir de l'échantillon de test, pour plus de détails se référer au papier de Richards et al. (2004).

### Coupe dans les diagrammes couleur-couleur

Pour rapprocher ces méthodes complexes des idées assez simples que l'on a des diagrammes couleur-couleur, Weinstein et al. (2004) fournissent les coupes dans l'espace des couleurs qui permettent de retrouver la plupart des quasars sélectionnés par les algorithmes du SDSS. La figure 3.17 montre les régions de l'espace des couleurs définies par l'intersection des coupes suivantes :

$$\begin{aligned}
 -0.07 &< u - g < 0.40 \\
 \{0.01 | -0.075\} &< g - r < 0.33 \\
 -0.20 &< r - i < 0.37 \\
 -0.19 &< i - z < \{0.215 | 0.29\} \\
 g - r &> 0.8333(u - g) - 0.2583 \\
 g - r &< 1.0625(r - i) + 0.3725 \\
 i - z &< -0.9730(r - i) + 0.4749
 \end{aligned}$$

avec  $\{x|y\}$  égal à  $x$  si  $r - i < 0.05$  et  $y$  si  $r - i \geq 0.05$ .

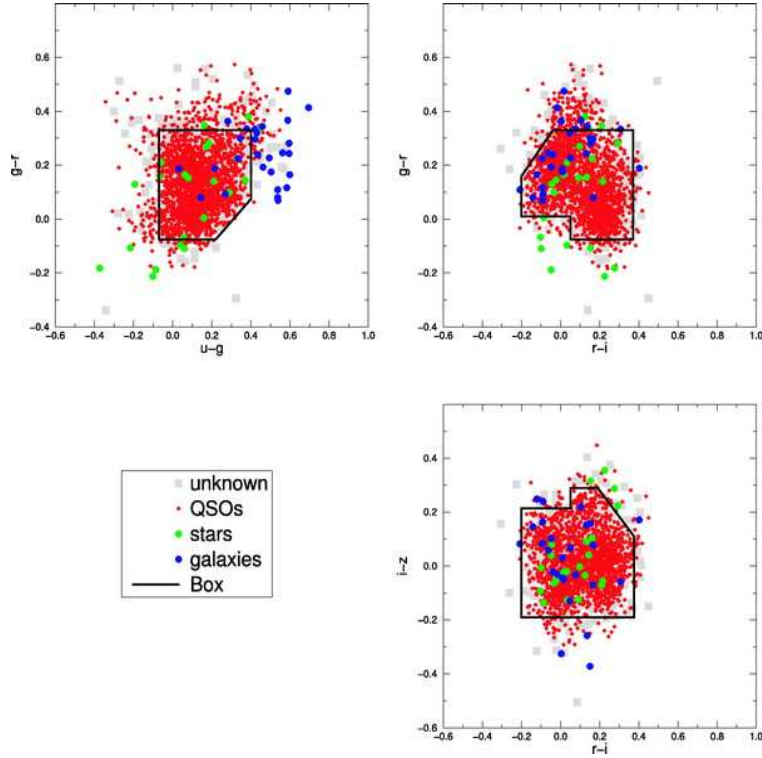


FIG. 3.17 – Diagramme couleur-couleur montrant les coupes en lignes noires pour sélectionner les quasars. Sont représentés les quasars avec un décalage spectral entre 0.8 et 2.2, ainsi que quelques contaminants. Issue de Weinstein et al. (2004).

### 3.3.5 Contaminants possibles à ce type de sélection

Parmi les contaminants les plus nombreux lors de cette sélection il faut essentiellement tenir compte de :

1. Les étoiles du halo qui ont une couleur bien plus bleue que celles du disque et constituent l'essentiel des contaminants lors de la sélection de quasars avec des décalages spectraux compris entre 2.5 et 3. Ceci est particulièrement vrai pour une population d'étoiles bleues, anciennes et évoluées, les étoiles bleues du halo 'Blue Horizontal Branch', soit étoiles BHB.
2. Un des contaminants les plus connus des sélections de quasars brillants à faibles décalages spectraux sont les naines blanches qui ont une couleur très bleue. Ces objets, issus de l'effondrement du noyau d'une étoile à la suite de l'arrêt du processus de fusion en son coeur se décomposent en plusieurs catégories selon les éléments qui dominent à leur surface. De façon générale il y a quatre variétés, les dA, dB, dO et dC (le 'd' signifiant dégénéré). A chacune sont associées des propriétés spectroscopiques liées à leur atmosphère : pour les types dA le spectre montre des raies d'hydrogène uniquement, l'atmosphère est constituée d'hydrogène pur. Les naines blanches dB n'ont que de l'hélium recouvrant leur noyau et présentent des raies d'hélium neutre, les dO sont aussi recouvertes d'hélium uniquement mais à plus grande température et on y observe donc des raies d'hélium ionisé. Les dC quant à elles ne montrent qu'un continu sans raies. Certaines naines blanches peuvent avoir un mélange d'éléments à leur surface, et leur dénomination est en conséquence : dAB pour des naines blanches contenant de l'hydrogène et de l'hélium neutre. Fan a effectué la simulation des naines blanche séparément de celle des étoiles, celle-ci s'organise de nouveau en quatre points :
  - 1) Tout d'abord il calcule le nombre de naines blanches à partir de la distribution spatiale (Widemann 1995) et de la fonction de luminosité (Liebert et al. 1988).
  - 2) Ensuite pour chacune de ces naines blanches sont déterminées aléatoirement leur luminosité et leur masse.
  - 3) Connaissant masse et luminosité, les magnitudes SDSS sont déterminées par des modèles d'atmosphère (Bergeron et al. 1995).
  - 4) Enfin l'extinction galactique et les erreurs photométriques sont ajoutées comme pour les étoiles.

Les résultats des simulations sont représentés figure 3.18

La contamination par les naines blanches lors de la sélection de quasars est finalement très limitée si l'on utilise plus de deux couleurs. En effet, ces contaminants se retrouvent essentiellement dans des sélections de quasars utilisant des excès  $U - V$ . Ces objets sont situés dans la partie la plus bleue de l'espace des couleurs, et c'est cette caractéristique qui permet de les séparer des quasars. En effet, si leur couleur  $u - g$  est proche de celle des quasars, les couleurs  $g - r$ ,  $r - i$ ,  $i - z$  sont situées dans les régions plus bleues que celles des quasars.



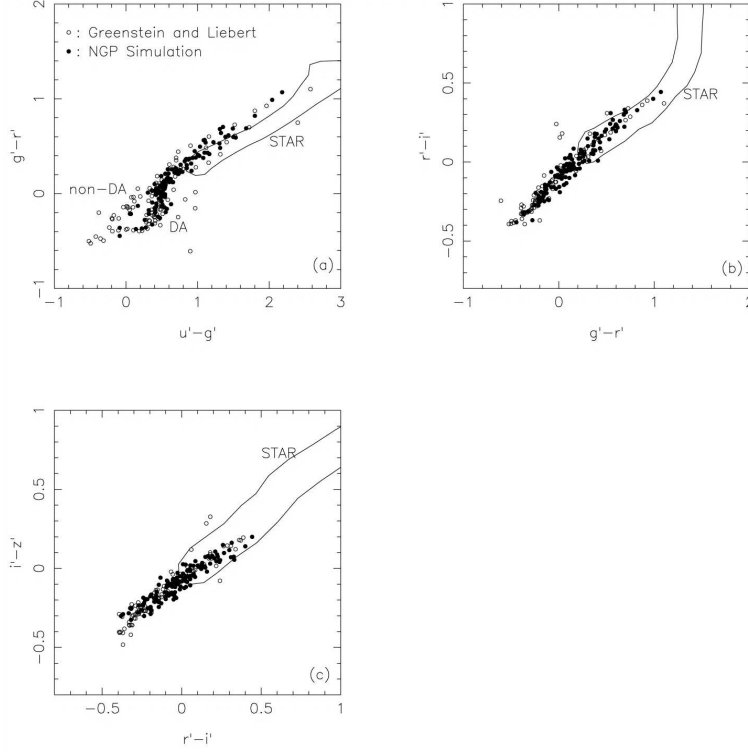


FIG. 3.18 – Diagramme couleur-couleur à partir d’une simulation de 20 degrés carrés avec  $g < 22.5$  (en cercles noirs). Les cercles blancs sont issus de l’atlas de Lenz et al. (1998). Les traits représentent les étoiles telles que simulées précédemment. Issue de Fan (1999).

La figure 3.19 résume les différents lieux des objets; il est important de se rappeler l’intersection des quasars avec les étoiles, qui seront les principaux contaminants.

J’ai décrit ici les derniers stades de l’utilisation des diagrammes couleur-couleur. Le principe de base reste le même, c’est-à-dire découper des régions dans l’espace des couleurs pour y sélectionner les objets. Les problèmes viennent tout simplement du découpage de ces régions. En effet, le découpage peut se faire dans un premier temps avec des simulations (Fan 1999). Mais au final l’algorithme de sélection doit utiliser les observations, et la découpe précise des régions nécessite un nombre important de données. Il faut donc d’abord avoir dépensé énormément de temps d’observation afin de réaliser une base de données conséquente. Mais la sélection a justement pour but de limiter le temps perdu à la confection de cette base de données. Passer de l’espace des couleurs du SDSS à celui du CFHT n’est pas évident et les données ne sont pas utilisables directement. Il est nécessaire de trouver une autre méthode qui permettrait de passer facilement d’un système photométrique à un autre.

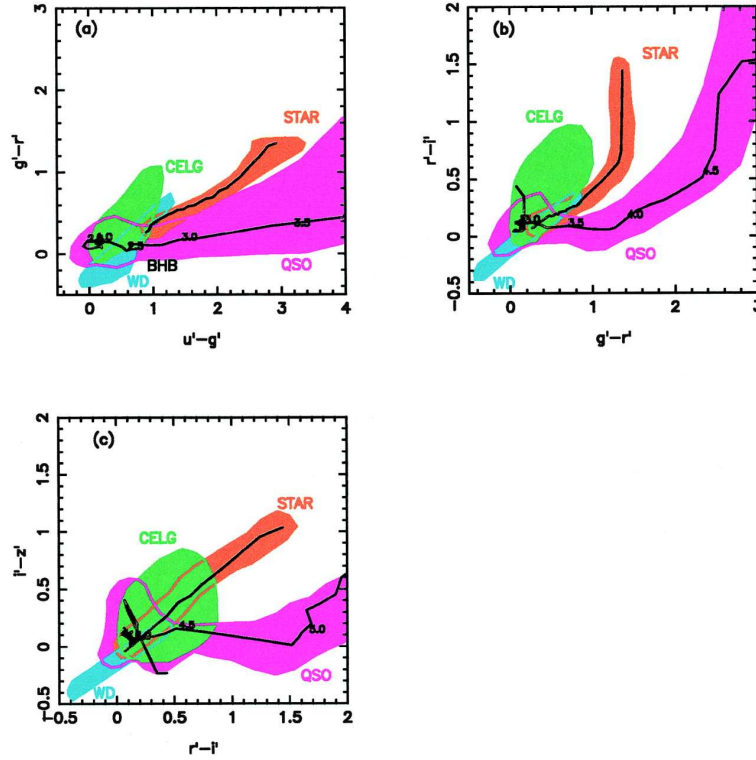


FIG. 3.19 – Résumé des différents lieux des objets et leurs intersections. CELG pour galaxie compacte à raie d'émission, BHB pour les étoiles bleues du halo et WD pour les naines blanches. Issue de Fan (1999).

### 3.4 Adaptation de distribution spectrale d'énergie

L'idée de comparer des observations photométriques à des spectres déjà connus d'objets est apparue lorsque l'on a voulu rechercher rapidement des types et des décalages spectraux de galaxies. En effet, le besoin était alors de mesurer un grand nombre de décalages spectraux de galaxies pour déterminer la valeur du paramètre de densité,  $\Omega_0$ . Baum (1962) a développé les prémices de la technique en comparant une distribution d'énergie spectrale (DSE) de galaxies aux décalages spectraux connus, à des observations photométriques d'autres galaxies elliptiques. Il a mesuré le déplacement entre les deux DSE et a pu estimer un décalage photométrique. La technique d'adaptation à des templates (ce terme désignant la connaissance a priori que l'on a des objets que l'on veut identifier, sous forme de spectres composites ou de simulations) a été développée par Loh & Spillar (1986). Les auteurs, pour tester leur méthode, ont observé à travers six filtres 34 galaxies dont le décalage spectral était connu, et l'ont ensuite appliquée à un champ contenant 1000 galaxies. Le principe de l'adaptation consiste donc à confronter des observations à un modèle, ceci peut se faire en calculant un  $\chi^2$  :

$$\chi^2 = \frac{1}{D} \sum_{i=1}^N \frac{[f_i^{obs} - s f_i^{templ}(z, T)]^2}{\sigma_i^2} \quad (3.12)$$

Avec  $N$  le nombre de filtres,  $D$  le nombre de degrés de liberté,  $\sigma_i$  est l'incertitude sur le flux dans le  $i^{eme}$  filtre et  $s$  est un facteur de normalisation qui réajuste le flux pour chaque template :

$$s = \frac{\sum_{i=1}^N f_i^{obs} f_i^{templ}(z, T) / \sigma_i^2}{\sum_{i=1}^N f_i^{templ}(z, T)^2 / \sigma_i^2} \quad (3.13)$$

Le flux des templates,  $f_i^{templ}(z, T)$ , dépendant à la fois du type de template utilisé,  $T$ , et du décalage spectral  $z$ .

### 3.4.1 Historique et principe

La première application de l'adaptation de distribution d'énergie spectrale à des données photométriques afin de sélectionner les quasars a été réalisée par Hatziminaoglou et al. (2000). L'idée étant de simplifier le problème qui consiste à utiliser l'ensemble des dimensions de l'espace des couleurs. Dans des diagrammes couleur-couleur, la sélection passe par un traitement multidimensionnel, tandis que l'adaptation de spectre connu remplace le problème par un traitement à une dimension, tout en utilisant l'ensemble de l'information de l'espace des couleurs. La technique permet de déterminer le meilleur modèle qui s'applique aux observations issues de relevés photométriques. Pour ce faire on peut par exemple utiliser l'outil HyperZ. La méthode permet à la fois de classer les objets et de leur attribuer un décalage spectral photométrique. Je reviendrai plus en détail sur HyperZ et son utilisation dans la section 4.1, je vais simplement présenter ici quelques résultats de Hatziminaoglou et al. (2000). Leur travail a surtout consisté à tester l'influence des jeux de filtres utilisés à travers des simulations. Les templates de quasars utilisés sont issus de spectres simulés, en changeant l'indice de la loi de puissance, sensée modéliser le continu, entre 0.0 et 1.0. Les raies d'émission sont issues de Peterson (1997). La forêt Lyman- $\alpha$  est modélisée selon Madau (1995) et aucun rougissement intrinsèque n'a été ajouté car ce paramètre est très mal maîtrisé pour les objets qui nous concernent. Les templates d'étoiles sont issus de Pickles (1998) plus 22 naines blanches et étoiles carbonées. L'identification des objets peut se faire en deux étapes : tout d'abord en choisissant le meilleur modèle entre étoile et quasar. Ensuite, à partir de cette sélection, exclure les objets dont la nature de quasar est rejetée à 95 ou 99 %, ce qui est appelé dans leur travail 'Sélection  $\chi^2$ '. Le nombre de degrés de liberté est égal dans ce cas au nombre de filtres moins 2 : un pour l'utilisation d'un filtre de référence, l'autre car c'est une procédure de minimisation. Ils définissent la complétude comme étant le rapport du nombre de vrais quasars parmi les candidats sur le nombre de quasars attendus ; le taux de confirmation comme étant le rapport du nombre de vrais quasars parmi les candidats sur le nombre total de candidats ; et enfin l'efficacité comme étant le produit de ces deux nombres. Les résultats portant sur les simulations sont résumés figure 3.20 et table 3.2.

Au final, le nombre de filtres n'est pas le paramètre le plus prépondérant, il est important de configurer la combinaison des filtres de façon à ce que chacun apporte des informations sur des parties différentes du spectre, comme la combinaison des filtres U et H/K. Néanmoins en dessous de 3 filtres la complétude ne

No	filtres	No	filtres	No	filtres	No	filtres	No	filtres
1	VRI	2	VRH	3	BRH	4	UBV	5	BRI
6	BVRI	7	BVR	8	VIK	9	BVRIH	10	BRIHK
11	VRIHK	12	BRHK	13	BVI	14	BRIK	15	VRIK
16	BVK	17	RIHK	18	BVRIHK	19	UBR	20	UBH
21	BVRIK	22	UBVR	23	UBRI	24	BVIK	25	UBVRI
26	UBVRIK	27	UBVRIH	28	UBRHK	29	UBHK	30	UBVRK
31	UVIK	32	UVRIK	33	UBVRIHK	34	UVK	35	UBRIHK
36	UBVIHK	37	UBVRHK	38	UVIHK	39	UVRIHK	40	UBRIK

TAB. 3.2 – Combinaisons des filtres utilisés dans la figure 3.20 classées par complétude croissante. Issue de Hatziminaoglou et al. (2000)

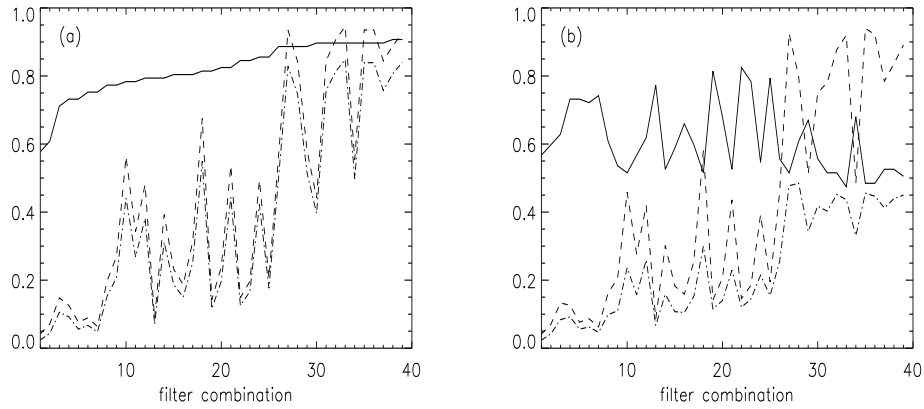


FIG. 3.20 – Complétude en trait solide, taux de confirmation en tirets et efficacité en tirets et points pour 40 combinaisons différentes de filtres. a) Sans la sélection  $\chi^2$ , b) avec la sélection à 99%. Issue de Hatziminaoglou et al. (2000)

varie pas malgré différentes combinaisons ; cela étant dû à la trop faible quantité d'information fournie par seulement 3 filtres. Pour ce qui est du type de sélection à employer, cela dépend de la stratégie du suivi spectroscopique, typiquement un taux de confirmation plus grand réduisant le temps d'observation (quoique cette influence dépend tout de même du type d'instrument utilisé). Enfin ils utilisent cet outil sur les champs du Deep Multicolor Survey (DMS, Osmer et al. 1998), du Hubble Deep Field (HDF) et le relevé de Parkes, ces tests sont résumés dans la table 3.3.

Les tests sur de vrais relevés ne sont faits que sur des petits échantillons dont les résultats sont présentés table 3.3. On y retrouve pour 3 champs (DMS, HDF et Parkes) leurs caractéristiques en terme de magnitude limite et de filtres utilisés, ainsi que les résultats sur les différents candidats quasars et étoiles avec les chiffres attendus. La sélection pour les quasars est faite de différentes façons, l'une avec la sélection dite du  $\chi^2$ , une autre sans, et aussi selon différents seuils de rejet. La surface couverte, et par conséquent le nombre d'objets sont très

Relevé	Surface	Mag. limite	Filtres	Etoiles		Candidats quasar				
				att	cand	att	95%	99%	no $\chi^2$	conf
DMS	0.83sd	B=22.3	U,B,V,R', I75,I86	3316	3332	106	295 (40)	348 (45)	388 (45)	51
HDF <sub>B</sub>	1.00sd	B=21.0	B,V,R	–		30	26	30	30	30
Parkes		H=19.6	B,V,R,I, J,H,K	–		129	99	107	125	129

TAB. 3.3 – La dernière colonne (conf) est le nombre de quasars qui ont été confirmés spectroscopiquement dans le relevé. Le nombre entre parenthèses est le nombre de quasars confirmés par spectroscopie qui sont parmi les candidats. Le nombre d'étoiles attendues vient de Robin et al. (1995), le nombre de quasars attendus de Hartwick & Schade (1990). Issue de Hatziminaoglou et al. (2000)

petits dans ces relevés, il sera donc important de voir comment se comporte l'outil sur de grandes statistiques; néanmoins on peut souligner l'une de leurs conclusions : “A realistic determination of the photometric errors is the key to a successful issue for this method”.

Une application de cette méthode est expliquée dans Wolf et al. (2001) (Object classification in Astronomical multi-colour survey). En fait ce qui est présenté n'est pas tout à fait une méthode d'adaptation de DSE, mais se situe plutôt entre celle-ci et les diagrammes couleur-couleur. En effet, l'algorithme proposé par Wolf et al. (2001) utilise une base de données de DSE (spectrale) pour chaque classe qu'ils désirent distinguer mais contrairement à ce qui a été fait par Hatziminaoglou et al. (2000), ils comparent les couleurs et non pas les flux des objets. De plus, ils mettent en place un 'kernel estimator' pour différencier les classes, et ce à partir du calcul d'une probabilité pour chaque objet d'appartenir à une classe, ce qui se rapproche de ce qui a été fait pour le SDSS. Cette méthode est donc une sorte de transition entre diagramme couleur et adaptation de DSE, j'ai choisi de la classer dans cette partie car selon les auteurs eux-mêmes celle-ci se rapproche plus des méthodes mettant en jeu les DSE. Je vais présenter rapidement la méthode utilisée, les bases de données pour chaque classe d'objets ainsi que les résultats sur différents relevés. La fonction de probabilité utilisée est séparée en trois classes :

$$p(c_i|\vec{q}) = \frac{p(\vec{q}|c_i)}{p(\vec{q}|c_{stars}) + p(\vec{q}|c_{galaxies}) + p(\vec{q}|c_{quasars})} . \quad (3.14)$$

où  $\vec{q}$  représente les caractéristiques de l'objet à classer, selon la valeur de cette fonction de probabilité ils décident à quelle classe l'objet va appartenir (citation de Wolf et al. 2001) :

If there is one class with  $p(c_i|\vec{q}) > 0.75$ , then we assign this class to

the object, but if all classes have likelihoods below 0.75, we call it *unclassifiable*.

Pour les objets qui ne rentrent pas dans ces critères, Wolf et al. regardent la distance dans l'espace des couleurs entre l'objet et chaque classe.

If an object is inconsistent at least at a confidence level of 99.73% (i.e.  $3\sigma$  in case of a normal distribution) with all members of all classes, then we call it *strange*.

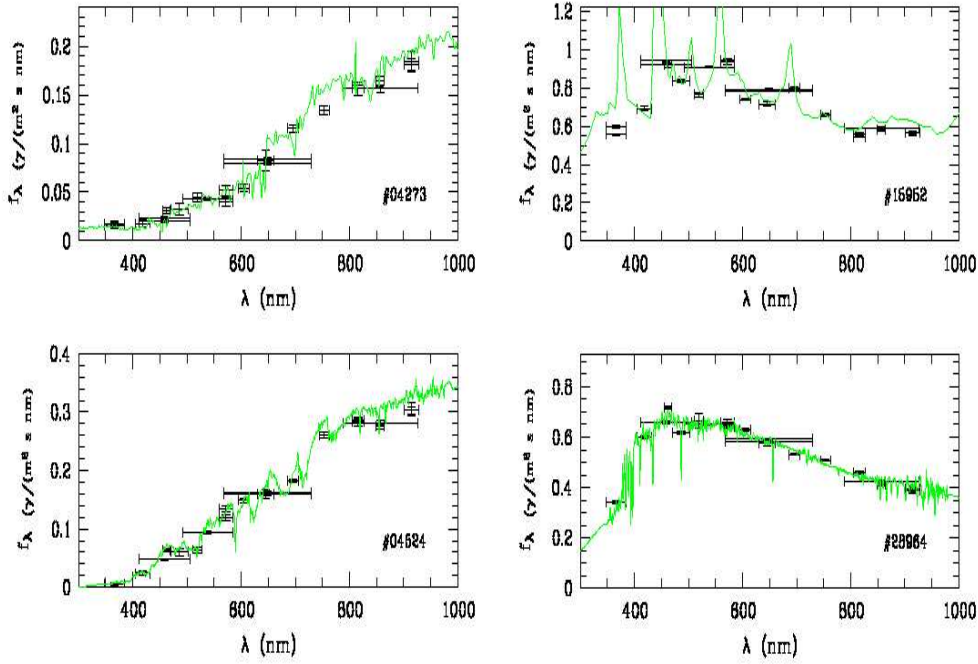


FIG. 3.21 – Utilisation des filtres moyennement étroits. Comparaison d'objets ayant des couleurs *BVRI* similaires, où les filtres de taille moyenne sont déterminants comparés aux filtres à large bande. En haut à gauche une galaxie à  $z=0.67$ , en bas une étoile de type M1. En haut à droite un quasar à  $z=2.6$ , en bas une étoile de type F2. Issue de Wolf et al. (2004).

L'équivalence des méthodes entre comparaison des flux et comparaison des couleurs vient de la façon dont la comparaison des flux se fait. En effet, Wolf et al. (2001) utilisent un filtre de référence où ils supposent connaître parfaitement les erreurs photométriques et où elles sont particulièrement faibles, ils définissent les couleurs à partir de ce filtre, et leur statistique revient, à l'approximation des erreurs près (cf. Wolf et al. 2001), à celle sur les flux.

Pour la définition des bases de données ils notent que celles-ci doivent couvrir tous les objets connus, jusqu'à une certaine limite. En effet, lorsque les modèles de comparaisons sont étendus jusqu'à couvrir même les objets les plus 'étranges' ou dont les caractéristiques s'éloignent franchement de la classe à laquelle ils appartiennent pourtant (par exemple certains BAL particulièrement forts pour les quasars), on introduit alors une confusion dans le classement des objets normaux. Ceci s'explique par le fait qu'en élargissant les différentes bases de

données des classes, elles finissent par se recouvrir, et la distinction devient donc plus complexe.

Leur choix de bibliothèque d'étoiles, identique à celui de Hatziminaoglou et al. (2000), s'est portée sur celle de Pickles (1998). La base de données de templates pour les quasars a évolué de la présentation de l'algorithme (Wolf et al. 2001) à son application dans Wolf et al. (2004) sur le relevé COMBO-17. Celui-ci permet la combinaison de 17 filtres, à bande large et moyenne. La dernière version utilise le composite déterminé par Vanden Berk et al. (2001) d'une part et d'autre part des spectres synthétiques ayant différents index de loi de puissance pour le continu et des raies d'émission de forces différentes. Ces templates synthétiques ont pour but de couvrir l'ensemble des DSE attendues de quasars. L'absorption dans la forêt Lyman- $\alpha$  est prise en compte par la fonction de Moller & Jakobsen (1990).

Pour les galaxies, ils utilisent les spectres synthétiques basés sur le code PEGASE (Fioc & Rocca-Volmerange 1997), ce qui change de leur précédente version où ils utilisaient les templates de Kinney et al. (1996).

Pour tester leur complétude ils utilisent le suivi spectroscopique de sources X dans le CDFS (The Chandra Deep Field South, Szokoly et al. 2004) ainsi que le 2QZ (spectroscopie dans les champs du 2dF Quasar Redshift Survey), et leur sélection permet de retrouver tous les quasars de ces champs, au nombre de 7. Ils ne précisent pas le nombre de candidats au total.

A noter aussi l'utilisation récente de l'adaptation de DSE pour déterminer les décalages spectraux dans Babbedge et al. (2004).

### 3.4.2 Comparaison aux diagrammes couleur-couleur

Il convient de s'interroger sur les différences entre les deux méthodes présentées, notamment sur les possibilités d'évolution. En ce qui concerne les diagrammes couleur-couleur, la base de données peut éventuellement s'enrichir d'un nombre plus important d'objets issus du relevé en cours permettant de mieux définir les régions d'exclusion. Mais ceci dépend forcément de l'efficacité de la méthode utilisée. En effet, la base de données ne peut s'enrichir que d'objets qui ont été sélectionnés par cette méthode, les améliorations sont donc limitées aux limites même de la sélection. A moins bien sûr de faire correspondre au relevés photométrique un relevé spectroscopique systématique. Dans le cas de l'adaptation de DSE, ce sont ces DSE qui peuvent être améliorées sans dépendre de l'état du relevé pour lequel elles sont utilisées.

Le traitement d'adaptation de DSE permet aussi d'estimer les caractéristiques des objets recherchés, essentiellement leur décalage spectral. Cette estimation peut être utilisée pour enrichir la sélection. Il est bien sûr possible par des méthodes mettant en jeu l'utilisation des couleurs de déterminer le décalage spectral (Weinstein et al. 2004), mais ceci se fait en réalisant un outil différent qui ne peut pas toujours être couplé à celui de la sélection.

Comme cela a déjà été montré par Hatziminaoglou et al. (2000), il est possible d'ajouter un traitement bayésien à la méthode d'adaptation si l'on a un fort a priori sur la population à sélectionner.

Le gros problème qui survient lors de l'utilisation de la méthode d'adaptation de DSE, c'est la nécessité de contraindre le mieux possible la photométrie. En effet, les résultats dépendent fortement de l'estimation des erreurs photométriques (voir en particulier Wolf et al. 2001). Cette dépendance s'accroît d'autant plus que l'on essaye d'élaborer des méthodes de sélection basées sur le calcul de probabilité, et sur des a priori de distribution. Une bonne partie des méthodes utilisant les diagrammes couleur-couleur n'utilisent pas les erreurs photométriques des objets, et s'ils l'utilisent cela correspond plus à l'estimation d'une distance entre deux objets, ce qui a un rôle moins important que dans le cas de calcul de  $\chi^2$ .

Les autres sources d'erreurs viennent de l'hypothèse selon laquelle on connaît parfaitement les réponses des instruments, filtres et détecteurs mis en jeu dans le relevé. Ceux-ci sont nécessaires pour passer de la base de données spectroscopique aux flux et couleurs des objets présents dans les catalogues d'observation. Wolf et al. (2001) proposent de comparer directement les couleurs ce qui permet de ne faire la transformation qu'une seule fois : de la base de données de modèles spectroscopiques à une base de données de couleurs.

Au final, la méthode d'adaptation de DSE semble plus adaptable et plus fiable d'un relevé à un autre ; elle se montre moins spécifique au relevé en présentant des résultats similaires. Il est possible de calibrer l'outil qui sera utilisé pour calculer les  $\chi^2$ , et ce à partir des DSE composites ou de l'avancée de nos connaissances des objets.





## Chapitre 4

# Méthodes de comparaison à des distributions spectrales d'énergie

*"there are many things we don't understand, and we just have to do the best we can with the knowledge we have"*

– Calvin

## Sommaire

---

<b>4.1</b>	<b>HyperZ - utilisation et qualification de l'outil. . . .</b>	<b>100</b>
4.1.1	Description du fonctionnement d'HyperZ . . . . .	100
4.1.2	Templates . . . . .	101
4.1.3	'Make catalogue' . . . . .	101
<b>4.2</b>	<b>Tests et sélection des candidats . . . . .</b>	<b>102</b>
4.2.1	Recherche de quasars . . . . .	102
4.2.2	Tests sur le SDSS . . . . .	104
	HyperZ et SDSS . . . . .	106
	Non détections . . . . .	108
	Absorption Lyman- $\alpha$ . . . . .	110
	Pas du décalage spectral . . . . .	112
	Choix des templates . . . . .	113
	Décalage spectral pour les tests . . . . .	118
	Critères de sélection . . . . .	119
	Nature des candidats . . . . .	123
4.2.3	Sélection sur le CFHTLS . . . . .	126
	Catalogues photométriques . . . . .	127
	Sélection . . . . .	128
	Suivi spectroscopique . . . . .	129
4.2.4	Conclusion . . . . .	130

---

Comment a été appliquée la méthode présentée dans le chapitre précédent, comment ai-je mené les tests sur le SDSS qui a été présenté chapitre 2 et enfin en quoi consiste ma sélection sur le relevé CFHTLS ? C'est ce que je vais décrire dans ce chapitre, en commençant par la présentation de l'outil utilisé : HyperZ.

## 4.1 HyperZ - utilisation et qualification de l'outil.

### 4.1.1 Description du fonctionnement d'HyperZ

La première étape réalisée par le code HyperZ (Bolzonella, Pello & Miralles, 2001 <http://webast.ast.obs-mip.fr/hyperz/>) est de lire les magnitudes observées avec les erreurs correspondantes et les magnitudes limites pour les objets non détectés. Il est possible de demander à HyperZ de calculer le dérougissement produit à l'intérieur des galaxies, ce qui n'a pas été utilisé lors de son utilisation dans mes tests car mal connu en ce qui concerne les objets qui nous concernent. Ensuite les magnitudes, qui peuvent être en système AB ou Vega, sont transformées en flux. Ceci se fait en passant de toute façon en système Vega, puisque le flux est calculé en se servant du spectre de Vega comme référence.

D'un autre côté vient la prise en compte des DSE, c'est-à-dire que le programme calcule les flux à partir des spectres de référence (i.e. des templates). Si l'on veut identifier des galaxies, il est possible de prendre en compte plusieurs paramètres tels que l'âge ou les types spectraux de galaxies. Chaque template est décalé d'un pas  $\Delta z$  jusqu'à un  $z_{max}$  et pour chaque pas en décalage spectral est appliquée l'absorption due à la forêt Lyman- $\alpha$ . Donc chaque template

d'entrée est transformé en  $\frac{z_{max}}{\Delta z}$  templates tenant compte de l'absorption due au milieu intergalactique (MIG). A ce niveau les spectres sont convolués avec les fonctions de réponse des filtres choisis, qui comprennent aussi la réponse du détecteur ; et ce afin d'obtenir les flux attendus. Une fois cette base de flux calculée, le programme procède au calcul de  $\chi^2$ , en considérant tous les décalages spectraux (selon les limites et le pas qui lui ont été fournis, il s'agit de la transformation des templates décrites précédemment), les types spectraux et l'âge. Par un procédé de minimisation le programme trouve le meilleur décalage spectral photométrique et la DSE correspondante. Il est possible d'obtenir en sortie d'HyperZ, le meilleur décalage spectral déterminé, la DSE observée, la meilleure DSE retenue et le  $\chi^2$ .

Un des principaux paramètres à prendre en compte est le catalogue photométrique d'entrée, contenant la photométrie des objets observés dans les différentes bandes du relevé. La photométrie doit être parfaitement connue, à travers les courbes de réponse des filtres utilisés pour le sondage ainsi que les erreurs associées aux magnitudes de chaque objet dans chaque bande. Il est clairement nécessaire de faire une première sélection sur les critères de qualité photométrique avant d'obtenir le catalogue définitif.

#### 4.1.2 Templates

Un élément central de cette méthode est la base des templates, aussi bien pour les étoiles que pour les quasars (si on se limite aux objets ponctuels, sinon il faut envisager des templates de galaxies). Il s'agit lors de la sélection/création de ces templates de tenir compte des différentes populations rencontrées et qui pourraient "polluer" l'identification finale. Par exemple il ne faut pas oublier d'inclure des spectres de naines blanches, car leurs caractéristiques photométriques sont proches de celles des quasars à bas  $z$ .

#### 4.1.3 'Make catalogue'

HyperZ comprend aussi une partie qui fonctionne en quelque sorte dans le sens inverse du code principal et qui est appelée 'Makecatalogue'. C'est-à-dire qu'à partir de DSE, des filtres et de quelques caractéristiques d'objets, il produit un catalogue photométrique. Makecatalogue ne tient pas compte de la répartition d'objet, il pioche de façon aléatoire dans les DSE qui lui sont fournis en entrée pour calculer le catalogue photométrique. Cet outil n'a donc pas été utilisé pour simuler un relevé en simulant la densité de tel objet dans telle direction du ciel.

Les catalogues que j'ai simulés pour effectuer les tests sont de deux sortes : l'un d'étoiles et l'autre de quasars. Pour les simulations d'étoiles j'ai systématiquement utilisé les DSE de Pickles (1998) avec aussi des spectres de naines blanches. La seule caractéristique que j'ai prise en compte pour les étoiles est l'éventuel décalage spectral de l'ordre de  $\pm 0.0002$  dû à leur vitesse propre. Pour les quasars j'ai soit pris une seule DSE : celle issue du composite du SDSS (Vanden Berk et al. 2001) et du HST (Telfer et al. 2002), soit un échantillon de DSE recouvrant différents indices spectraux pour les lois de puissances (ceux décrits

dans la section 3.4.1). De plus, j'ai fait varier les décalages spectraux entre 0 et 7 avec un pas de 0.001. MakeCatalogue tient compte de l'absorption par la forêt Lyman- $\alpha$  selon Madau (1995).

Lors de la production du catalogue photométrique il est possible d'inclure différents types d'erreur associées aux magnitudes : soit une erreur fonction de la magnitude de l'objet telle que :

$$\Delta m(m) \simeq \sqrt{[2.5 \log(1 + \frac{1}{S/N})]^2 + ERR\_MAG\_MIN^2} \quad (4.1)$$

et  $S/N$  qui est le rapport signal sur bruit est donné par la magnitude :

$$S/N = (S/N)_0 10^{-0.4(m-m_0)} \quad (4.2)$$

avec  $(S/N)_0$  le rapport signal sur bruit à la magnitude de référence  $m_0$  ; ces deux paramètres sont donnés en entrée du programme ainsi que l'erreur minimale à considérer ( $ERR\_MAG\_MIN$ ). Dans le second cas d'erreur, il est possible d'appliquer un bruit aléatoire avec un écart-type en magnitude fixé, qui est donné en entrée du programme. L'application d'erreur sur les données simulées nous permettra de tester le comportement de l'outil dans les cas où l'on utilise des données avec des erreurs photométriques importantes, ce qui est le cas lorsque l'on utilise des relevés de plus en plus profonds en magnitude.

## 4.2 Tests et sélection des candidats

J'en viens désormais à la partie où il s'agit de sélectionner les candidats quasars. C'est-à-dire définir les critères qui vont me permettre de classer les objets. Pour définir ceux-ci j'ai effectué plusieurs tests avec les données SDSS. En effet, cette base de données, proche des caractéristiques du CFHTLS, permet d'effectuer des tests sur des échantillons significatifs et de rencontrer nombre de cas de figure. Je vais dans un premier temps présenter les tests effectués avec le SDSS et ensuite comment il a fallu adapter les critères retenus aux données issues du CFHTLS.

### 4.2.1 Recherche de quasars

Les différentes recherches de quasars qui ont déjà eu lieu permettent de connaître la densité des quasars en fonction de la profondeur en magnitude à laquelle on va les rechercher. Vanden Berk et al. (2001) proposent une relation du nombre de quasars en fonction de la magnitude en  $i$  :

$$N_{<i} = 10^{0.738i-13.08} \text{ deg}^{-2} \quad (4.3)$$

$N_{<i}$  est le nombre de quasars par degré carré avec une magnitude en  $i$  inférieure à la limite spécifiée. Leur travail a porté sur la complétude du SDSS pour des magnitudes  $i < 19.1$ , cette relation n'est donc à considérer que pour cet ordre de grandeur en  $i$ . Une étude plus récente de Richards et al. (2006) , présentent une évolution de la densité des quasars en fonction de la magnitude dans le filtre  $g$  du

SDSS. Il faut garder à l'esprit que pour le SDSS la magnitude limite de sélection a été choisie dans le filtre  $i$ . La figure 4.2 et la table 4.1 rendent compte à la fois de l'étude menée par Richards et al. (2006) et rappellent aussi les chiffres obtenus avec le 2QZ. Cette dernière table comprend le nombre de quasars par 0.25 magnitudes et par degré carrés ( $N(g)$  en filtre  $g$  et  $N(i)$  en filtre  $i$ ), le nombre de quasar par degré carré plus brillants que la magnitude indiquée ( $N(< i)$ ).  $N_Q$  est le nombre de quasars observés dans chaque intervalles de 0.25 magnitudes sur une surface effective de 1622 degrés carrés. La distribution en décalage spectral des quasars du SDSS est représentée figure 4.1. Une estimation à des magnitudes plus profondes provenant des relevés 2dF et 2SLAQ est présenté table 4.2. Au final on peut s'attendre à trouver entre 150 et 300 quasars par degré carré en se limitant à une magnitude de 22.5 en  $g$ .

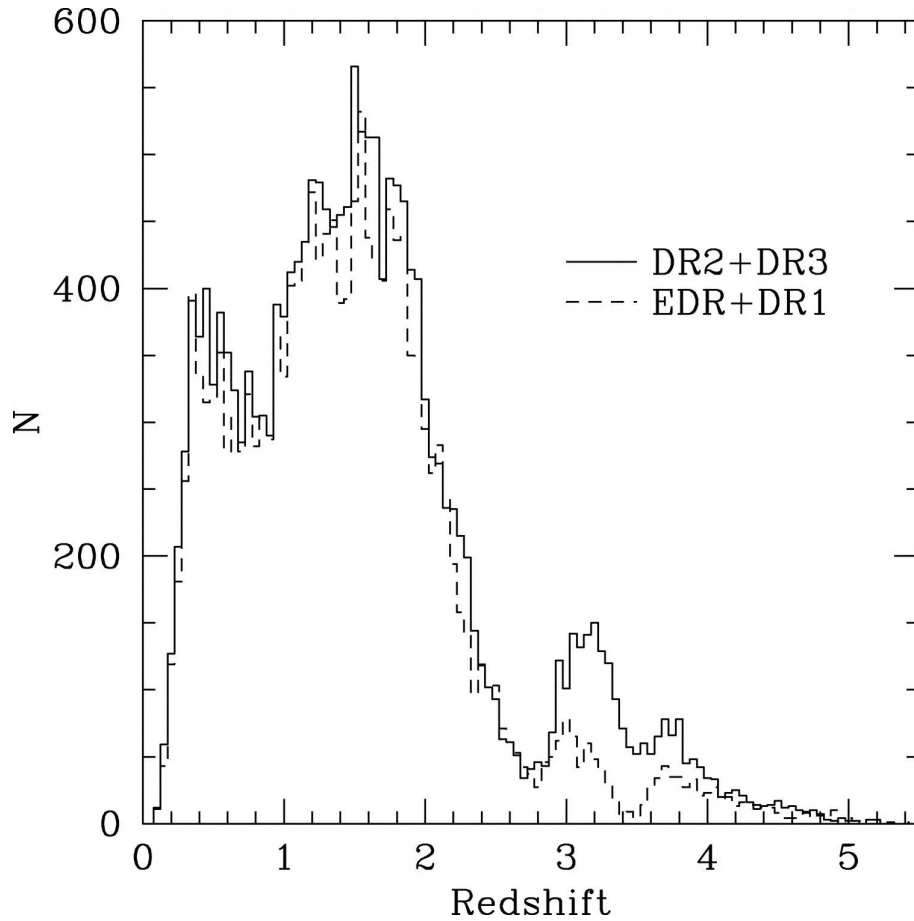


FIG. 4.1 – Comparaison de l'évolution de la distribution des quasars en fonction du décalage spectral selon les différentes évolutions du relevé. Les structures observées dans la distribution en décalage spectral sont dues aux effets de sélection (Richards et al. 2006). Issue de Richards et al. 2006

mag	N( <i>g</i> )	N(< <i>g</i> )	N <sub>Q</sub>	N( <i>i</i> )	N(< <i>i</i> )	N <sub>Q</sub>
15.475	0.00 ± 0.00	0.00 ± 0.00	2	0.00 ± 0.00	0.00 ± 0.00	1
15.725	0.00 ± 0.00	0.00 ± 0.00	3	0.00 ± 0.00	0.00 ± 0.00	3
15.975	0.00 ± 0.00	0.01 ± 0.00	4	0.00 ± 0.00	0.00 ± 0.00	2
16.225	0.01 ± 0.00	0.01 ± 0.00	11	0.01 ± 0.00	0.01 ± 0.00	10
16.475	0.01 ± 0.00	0.02 ± 0.00	10	0.01 ± 0.00	0.02 ± 0.00	12
16.725	0.02 ± 0.00	0.04 ± 0.00	32	0.01 ± 0.00	0.03 ± 0.00	23
16.975	0.06 ± 0.01	0.10 ± 0.01	85	0.04 ± 0.01	0.08 ± 0.01	67
17.225	0.08 ± 0.01	0.17 ± 0.01	117	0.07 ± 0.01	0.15 ± 0.01	107
17.475	0.12 ± 0.01	0.29 ± 0.01	190	0.11 ± 0.01	0.26 ± 0.01	174
17.725	0.23 ± 0.01	0.53 ± 0.01	357	0.21 ± 0.01	0.47 ± 0.01	327
17.975	0.31 ± 0.01	0.83 ± 0.01	472	0.33 ± 0.01	0.80 ± 0.01	511
18.225	0.56 ± 0.02	1.40 ± 0.02	868	0.55 ± 0.02	1.35 ± 0.02	849
18.475	0.78 ± 0.02	2.18 ± 0.02	1202	0.82 ± 0.02	2.17 ± 0.02	1257
18.725	1.19 ± 0.03	3.37 ± 0.03	1839	1.25 ± 0.03	3.42 ± 0.03	1923
18.975	1.70 ± 0.03	5.07 ± 0.03	2620	1.86 ± 0.04	5.28 ± 0.03	2870
19.225	1.61 ± 0.03	6.68 ± 0.03	2484	2.62 ± 0.04	7.90 ± 0.04	4028
19.475	0.56 ± 0.02	7.24 ± 0.02	855	...	...	900
19.725	0.13 ± 0.01	7.37 ± 0.01	206	...	...	216
19.975	0.03 ± 0.01	7.41 ± 0.00	53	...	...	55
20.225	0.01 ± 0.00	7.42 ± 0.00	19	...	...	20

TAB. 4.1 – Densité de quasars en fonction de la magnitude ( $0.3 < z < 2.2$ ). Voir le texte pour les légendes. Issue de Richards et al. (2006)

mag ( <i>g</i> )	B00	C04	2SLAQ
20.0	15.87	17.50	18.96
20.5	26.99	28.27	31.09
21.0	41.68	40.22	47.79
21.5	59.45	52.01	69.13
22.0	78.88	62.46	93.77

TAB. 4.2 – Densité de quasars à faibles magnitudes. Comparaison de l'estimation des densités de quasars en fonction de la magnitude ( $0.3 < z < 2.2$ ). B00 correspond à Boyle et al. (2000), C04 à Croom et al. (2004) et la dernière colonne à une estimation provenant du 2SLAQ (la limite du relevé étant  $g = 21.85$ ). Les unités sont par degré carré. Issue de Richards et al. (2005)

#### 4.2.2 Tests sur le SDSS

La photométrie du SDSS est mesurée non pas en magnitude conventionnelle du type AB ou Vega mais en ce qu'ils appellent des 'asinh magnitudes' (appelées aussi 'luptitudes' cf. paragraphe 2.1). Ors HyperZ ne traite que des magnitudes de type AB ou Vega, et effectue la conversion du système AB à Vega. (cf. la partie 4.1). Il a fallu donc dans un premier temps convertir ces luptitudes en magnitude AB :

$$m_{AB} = -2.5 \cdot \log \left[ 2b \cdot \sinh \left( \frac{-m_s \log(10)}{1.08574} - \frac{2.5}{1.08574} \log(b) \right) \right] \quad (4.4)$$

(à partir de l'équation 2.7 présentant les magnitudes asinh). Les valeurs du paramètre  $b$  sont dans la table 2.1. De même les erreurs estimées sont :

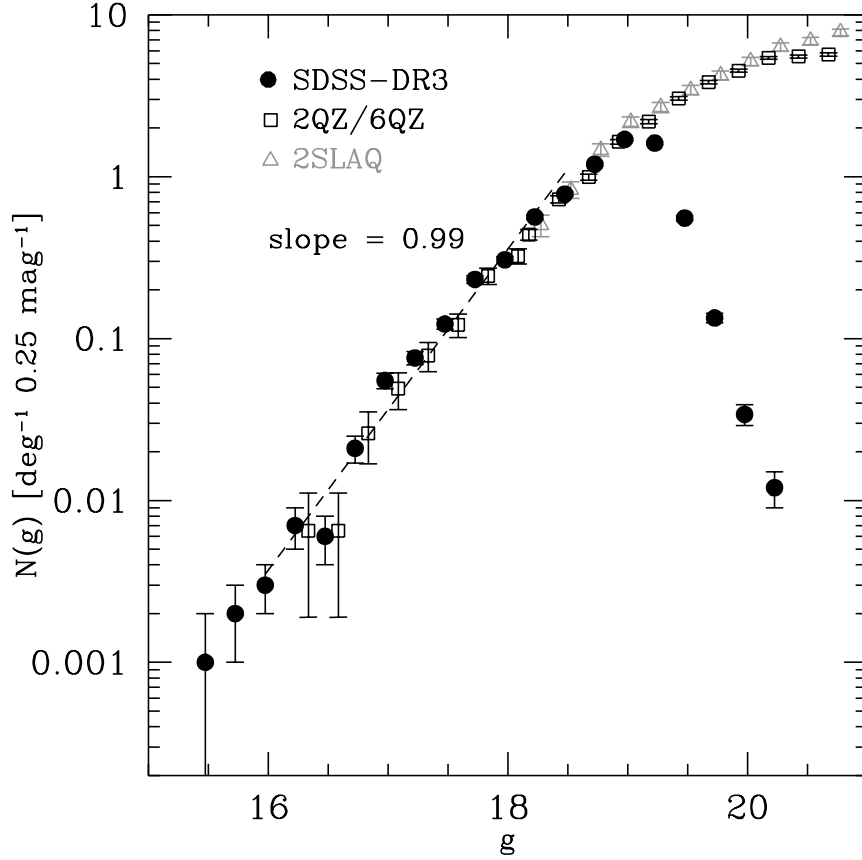


FIG. 4.2 – Nombre de quasar sélectionnés dans le SDSS en cercles noirs. Les données du 2QZ/6QZ (Croom et al. 2004) sont données par les carrés et les données du 2SLAQ (Richards et al. 2005) par les triangles. La chute des données du SDSS est due à l'effet de sélection en magnitude limite en  $i$ . Issue de Richards et al. (2006)

$$\Delta m_{AB} = \Delta m_s \left( \frac{1}{\left( \sinh\left(-\frac{m_s \log(10)}{1.08574} - \frac{2.5}{1.08574} \log(b)\right)\right)^2} + 1 \right) \quad (4.5)$$

Les luptitudes ne sont pas très différentes des magnitudes de type AB. La différence ne se voit qu'à très faibles magnitudes (cf. figure 2.2). Pour rappel les magnitudes Vega s'écrivent :

$$m_{Vega} = -2.5 \cdot [\log\left(\int R(\lambda) f(\lambda) d\lambda\right) - \log\left(\int R(\lambda) f_{Vega}(\lambda) d\lambda\right)] \quad (4.6)$$

avec  $f$  le flux de l'objet dont on mesure la magnitude et  $f_{Vega}$  le flux de Vega ;  $R$  est la réponse instrumentale.



## HyperZ et SDSS

Afin de clarifier le texte, il est important de se souvenir que jusqu'à la partie sur les critères de sélection, lorsque je parlerai de candidats, ceux-ci auront été sélectionnés par HyperZ sur le critère :  $\chi_{qso}^2 < \chi_{star}^2$ . Il s'agit d'un critère très simple et naturel qui va me permettre de présenter les autres effets à prendre en compte et qui se situent en amont de la sélection.

Dans un premier temps j'ai voulu tester la stabilité d'HyperZ et éprouver son fonctionnement dans un environnement parfaitement contrôlé. Pour cela j'ai utilisé l'outil Makecatalogue décrit précédemment (section 4.1.3). Celui-ci m'a permis de réaliser des simulations de catalogues photométriques correspondants à ceux du SDSS. Bien sûr ceci est valable jusqu'à une certaine extension, c'est-à-dire que je ne prends pas en compte la zone de pointage du ciel ni la densité des objets (cf . section 4.1.3). Il s'agit dans ces simulations de reproduire les couleurs d'objets à partir des spectres connus (DSE) et des filtres du relevé, tenant compte à la fois de la réponse des filtres et du détecteur (et de l'atmosphère dans le cas du SDSS). Il s'agit du même type de simulation que celle que j'ai utilisée pour comparer les observations aux données de Pickles (cf . section 2.1).

Les simulations peuvent inclure des erreurs de différents types. Dans un premier temps, j'ai voulu estimer l'impact de l'évaluation des erreurs sur la capacité d'HyperZ à affecter un décalage spectral photométrique correct, et ce sur une population de quasars observés avec les filtres du SDSS. Les templates utilisés devaient être les plus simples possibles, il ne s'agit pas dans ces tests d'éprouver la reconnaissance de différents templates, mais seulement l'influence des erreurs et la robustesse d'HyperZ. J'ai donc pris un spectre composite issu d'une combinaison entre SDSS (Van denberk et al. 2001) et HST (Telfer et al. 2002) et les 5 filtres du SDSS. Dans tous ces cas de figure, HyperZ identifie correctement les quasars, et ce à plus de 98%, avec un critère simple qui compare les  $\chi^2$  associés aux modèles d'étoiles à ceux associés aux modèles de quasar.

Les résultats des tests sont illustrés dans la figure 4.3. Sont représentés le décalage spectroscopique déterminé par HyperZ en noir ainsi que les valeurs des  $\chi^2$  (en vert) associés aux quasars en fonction du décalage réel (issu des modèles). Dans trois cas, le panneau du haut avec des erreurs moyennes lors des simulations de 0.005 (NOISE\_TYPE=1 et SN=10 dans HyperZ), de 0.02 pour celui du milieu et de 0.1 en bas. Lorsque les erreurs sont parfaitement maîtrisées les décalages spectroscopiques sont dans la plupart des cas quantitativement retrouvés, les dégénérescences étant directement liées aux domaines couverts par les filtres. L'information pour des quasars à faible décalage spectral est ambiguë, mais dès que l'on rentre dans le domaine de longueur d'onde où l'émission Lyman- $\alpha$  rentre dans le domaine couvert par le filtre  $u$ , le décalage spectral est plus facile à déterminer (pour des décalages spectraux supérieur à  $z \sim 2$ ). Quelques dégénérescences subsistent, essentiellement à cause du faible recouvrement entre les filtres. L'augmentation des erreurs a pour effet d'augmenter l'incertitude sur les décalages spectraux déterminés et sur la dispersion des valeurs de  $\chi^2$  calculés (l'échelle est arbitraire pour les  $\chi^2$ ).

Il est clair qu'avec des erreurs sur les magnitudes inférieures à 0.1 la dé-

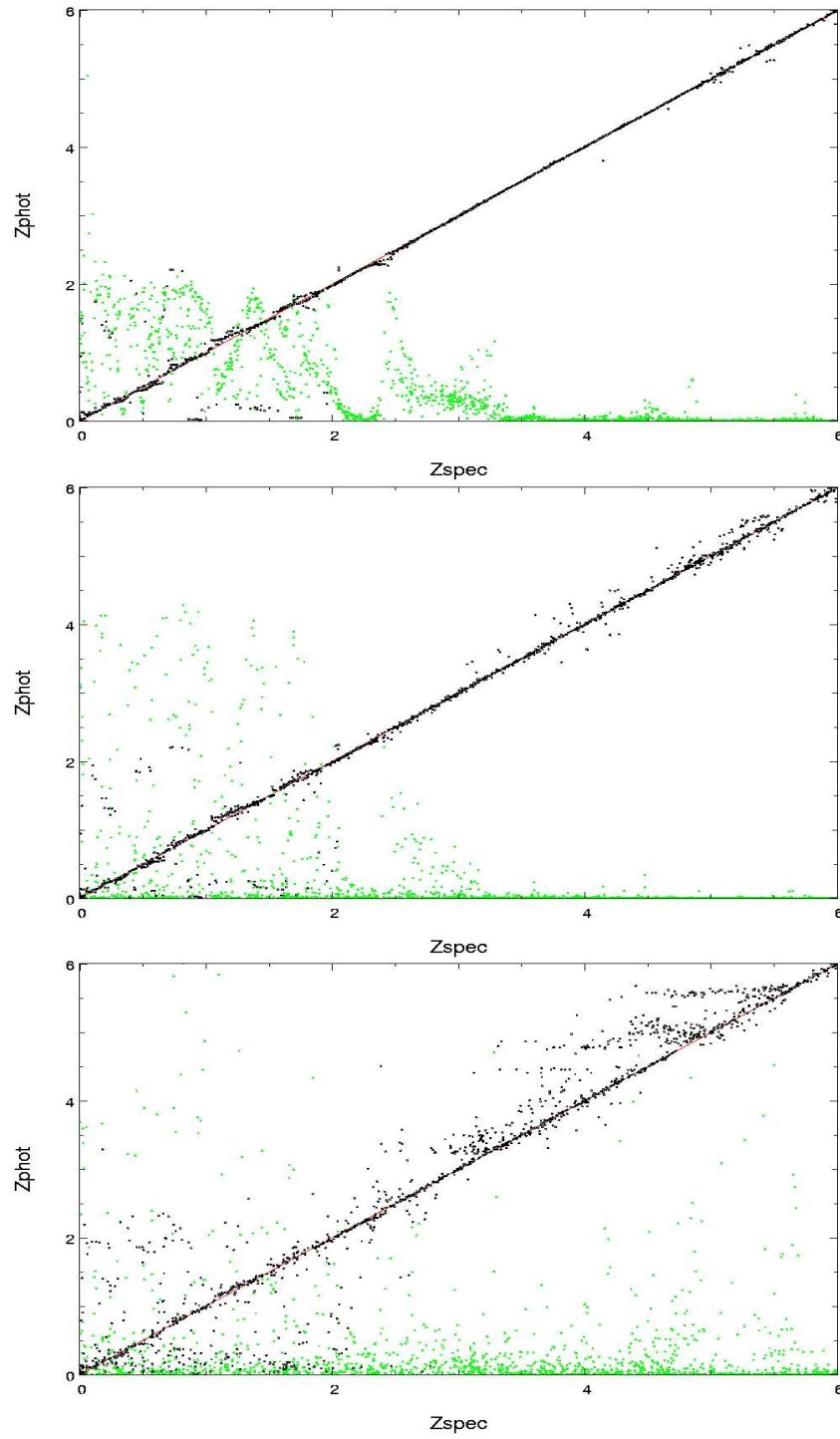


FIG. 4.3 – Graphique représentant le décalage spectroscopique déterminé par HyperZ en noir et les valeurs des  $\chi^2$  (en vert) associés aux quasars en fonction du décalage réel (issu des modèles). Panneau du haut avec une erreur moyenne de 0.005, du milieu avec une erreur moyenne de 0.02, panneau du bas avec une erreur moyenne de 0.1.

termination du décalage spectral et l'identification des objets est très aisée et extrêmement fiable. Au delà de 0.5, le décalage spectral n'est plus fiable à 0.3 près. L'identification, et surtout les valeurs des  $\chi^2$  associés aux quasars évoluent énormément, mais ceci est attendu par la définition même des  $\chi^2$ . L'évaluation précise des erreurs joue un rôle essentiel si l'on veut effectuer une sélection en se basant sur des critères tenant compte des valeurs des  $\chi^2$ , comme le montre les valeurs des  $\chi^2$  calculés en fonction de chaque type d'erreur. Au final HyperZ, et c'est bien normal, est sensible à l'évaluation des erreurs photométrique associées à chacun des objets pour déterminer le décalage spectral mais semble robuste en ce qui concerne l'identification.

### Non détections

La photométrie adoptée par le SDSS a ceci de particulier qu'il n'y a pas en soit de seuil de non-détection, mais qu'il faut le définir. En effet, le fait d'utiliser une fonction de sinus hyperbolique permet d'associer une magnitude même si le flux est à la limite. Mais lors de la transformation en magnitude de type AB qui est nécessaire pour le traitement par HyperZ, j'ai dû appliquer une valeur de non détection lorsque le flux était en deçà du flux limite. J'ai simplement appliqué la même convention que pour SourceExtractor (Bertin & Arnouts 1996) à savoir appliquer une valeur de 99 au catalogue photométrique lorsque la magnitude ainsi transformée tombe en dessous de la magnitude limite pour chaque filtre (cf. tableau 2.1). C'est ce type de catalogue dont je disposerai pour les relevés issus du CFHTLS.

Ceci fait, il faut encore déterminer quel comportement adopter face à une non-détection, en sachant que la plupart des non-détections auront lieu dans le filtre  $u$  pour les quasars. HyperZ propose quatre possibilités :

1. Le filtre n'est pas pris en compte lors du calcul du  $\chi^2$ . Ceci permet d'éviter d'avoir trop de degrés de liberté lors de la minimisation mais enlève une partie de l'information, car l'absence de détection nous donne tout de même une magnitude limite.
2. Le flux dans ce filtre est mis à zéro avec une erreur correspondante au flux limite de détection.
3. Le flux et l'erreur sont tout deux fixés à la moitié du flux limite.
4. Le flux et l'erreur associée sont ceux du flux limite et de l'erreur qui y est associée.

Pour vérifier simplement l'influence du nombre de degrés de liberté sur le fonctionnement d'HyperZ, et plus particulièrement sur la détermination du  $\chi^2$ , j'ai comparé les deux premières méthodes via des simulations. Les simulations sont identiques à celles décrites précédemment, avec une erreur moyenne de 0.02. Ces simulations ont été utilisées en entrée d'HyperZ avec deux paramétrages différents pour les filtres (cas 1) puis 2) décrit précédemment). J'ai voulu comparer l'évolution des  $\chi^2$  en fonction des décalages photométriques dans le premier et le second cas. (cf. figure 4.4).

Il est clair que pour la détermination des décalages spectraux et l'identification des objets les deux cas sont très similaires, même si la tendance des

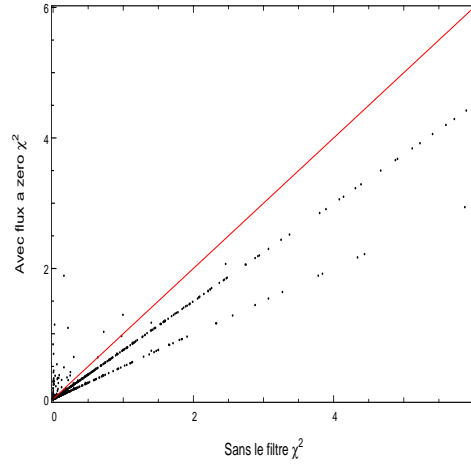


FIG. 4.4 – Graphique représentant les valeurs de  $\chi^2$  (les points noirs) évalués en mettant le flux à zéro (cas 2) en fonction des  $\chi^2$  évalués sans tenir compte du filtre où il n'y a pas eu de détection (cas 1). La courbe rouge représente la limite où les deux  $\chi^2$  sont égaux.

$\chi^2$  montrent que ceux-ci ont des valeurs bien inférieures pour le second cas, qui semble donc meilleur. De plus si un objet n'est détecté que dans un filtre, l'utilisation de la première option empêche le traitement de l'objet, il est donc préférable d'utiliser un flux à zéro avec une erreur correspondante au flux limite de détection.

Pour ce qui est des autres possibilités, j'ai préféré juger à partir des données, car les simulations ne peuvent pas tout à fait reproduire les effets de non-détections liées aux limites en magnitude des relevés. J'ai effectué ce test pour les quasars qui n'ont pas été détectés dans au moins un des filtres, issus du catalogue de Schneider et al. (2003, 2005). Je présente les résultats pour Schneider (2005). Je mets ce catalogue photométrique en entrée d'HyperZ (il contient des objets qui ne sont pas détectés dans certains filtres) et je récupère la valeur de  $\chi^2$  qui est attribué à chaque objet lorsqu'ils sont comparés à des DSE de quasars. Dans chacun des cas j'ai cherché à tracer la même évolution des  $\chi^2$  que celle vue avec les simulations (figure 4.4) et j'ai aussi voulu vérifier la capacité d'HyperZ à identifier correctement les quasars (c'est-à-dire attribuer aux objets une DSE de quasar plutôt que celle d'une étoile). Dans la comparaison des cas 2), 3) et 4), les cas 3) et 4) donnent des valeurs de  $\chi^2$  légèrement supérieures (3.8 et 4 en moyenne au lieu de 3.6 pour le second choix, cf. figure 4.5). De plus parmi les trois derniers choix, l'identification d'objets apporte le même pourcentage de quasar sélectionnés (i.e. une complétude de 90%).

En conclusion, afin de ne pas trop dépendre des caractéristiques des différents relevés et rester aussi prudent que possible, j'ai opté pour la seconde option dans le paramétrage d'HyperZ. En effet, une non détection apporte des informations, essentiellement lorsqu'il s'agit des quasars à grand décalage spectral qui n'ont pas ou très peu de flux en bande  $u$  à cause de la présence de la forêt Lyman- $\alpha$  dans ce domaine. De plus, d'après les valeurs de  $\chi^2$ , c'est cette option qui permet de mieux adapter les modèles aux observations.

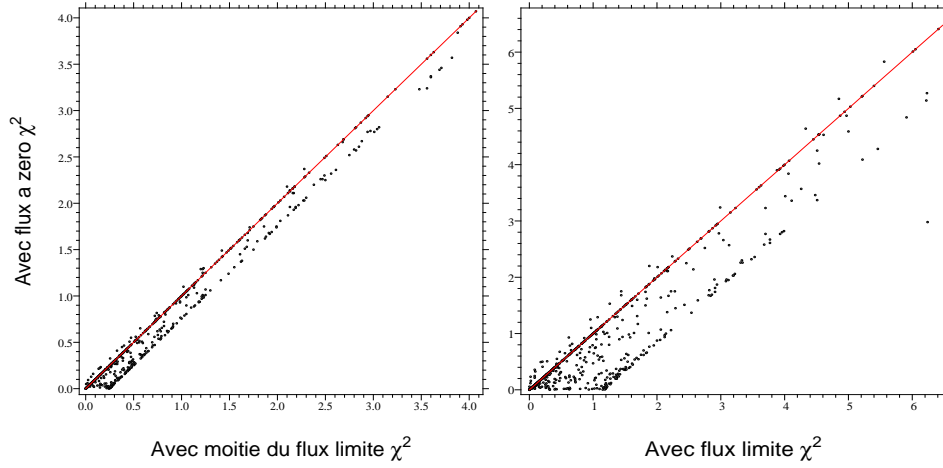


FIG. 4.5 – Graphique représentant les valeurs de  $\chi^2$  calculés en mettant le flux à zéro en fonction des  $\chi^2$  calculés en utilisant la moitié du flux limite. Même figure dans le panneau de droite, avec le flux limite et son erreur associée en abscisse. La plupart des points qui montrent des valeurs de  $\chi^2$  égales (sur la ligne rouge) correspondent à des quasars à grand décalage spectral ( $z > \sim 4$ ) dont l'identification est aisée.

### Absorption Lyman- $\alpha$

Dans les parties précédentes j'ai décrit comment le milieu intergalactique est traversé par la lumière des quasars et laisse son empreinte dans leurs spectres, ce qui au final influe sur les couleurs observées des quasars. Afin de tenir compte de cette influence pour la reconnaissance des objets, il faut pouvoir modéliser cette absorption, en l'incluant dans les distributions spectrales d'énergie des quasars. Je vais reprendre ici ce qui est utilisé au sein d'HyperZ (et qui est décrit dans le manuel) et ce qui a été testé.

Cette absorption se situe en-deçà de l'émission Lyman- $\alpha$  du quasar (1216 Å au laboratoire). On considère deux paramètres caractérisant la quantité d'absorption (Oke & Korycansky 1982), d'une part celle entre Lyman- $\alpha$  et Lyman- $\beta$  (en-deçà de 1026 Å, longueur d'onde du laboratoire) et d'autre part celle entre Lyman- $\beta$  et la limite de Lyman. Les régions concernées dans les spectres des quasars sont celles de la forêt Lyman- $\alpha$  et des suivantes (Lyman- $\beta$ ,  $\gamma$ ,  $\delta$  ...), bien sûr ces deux forêts se superposent dans le spectre en-deçà de la longueur d'onde de la transition Lyman- $\beta$  au décalage spectral du quasar émetteur. Pour modéliser cette absorption, Oke & Korycansky (1982) ont observé des quasars à grand décalage spectral et ont défini deux paramètres pour les absorptions : DA entre l'émission Lyman- $\alpha$  et Lyman- $\beta$ , et DB entre l'émission Lyman- $\beta$  et la limite de Lyman. HyperZ utilise les facteurs calculés par Madau (1995), qui dépendent du décalage spectral :

$$\langle D_A \rangle = 1 - \frac{1}{\Delta\lambda_A} \int_{1050(1+z_{em})}^{1170(1+z_{em})} \exp(-\tau_{eff}^\alpha) d\lambda_{obs} \quad (4.7)$$

$$\langle D_B \rangle = 1 - \frac{1}{\Delta\lambda_B} \int_{920(1+z_{em})}^{1015(1+z_{em})} \exp(-\tau_{eff}^\beta) d\lambda_{obs} \quad (4.8)$$

$$(4.9)$$

La profondeur optique pour chacune des forêts est déterminée comme :

$$\tau_{eff}^\alpha = A_2 \left( \frac{\lambda_{obs}}{\lambda_\alpha} \right)^{3.46}; \tau_{eff}^\beta = \sum_{j=3,11} A_j \left( \frac{\lambda_{obs}}{\lambda_j} \right)^{3.46} \quad (4.10)$$

avec  $\Delta\lambda_A = 120(1+z_{em})\text{\AA}$  et  $\Delta\lambda_B = 95(1+z_{em})\text{\AA}$ .  $A_2$  correspond à la contribution de la forêt Lyman- $\alpha$  et est égal à  $3.6 \cdot 10^{-3}$ . Les contributions à  $\langle D_B \rangle$  viennent des séries Lyman, beta, gamma et delta avec les coefficients :  $A_3 = 1.7 \cdot 10^{-3}$ ,  $A_4 = 1.2 \cdot 10^{-3}$  et  $A_5 = 9.3 \cdot 10^{-4}$ .

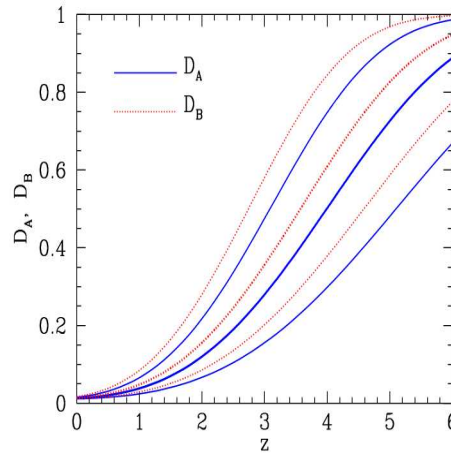


FIG. 4.6 – Graphique représentant les valeurs de  $\chi^2$  d'une part en utilisant la moitié du flux limite (bleu) et en mettant le flux à zéro (rouge), en fonction du décalage spectroscopique. Issue du manuel d'HyperZ (Bolzonella, Pello, Miralles 2001)

L'absorption due au HI peut varier significativement d'une ligne de visée à une autre. Toutefois il faut éviter de laisser trop de degrés de liberté à l'outil lors de la procédure de minimisation. J'ai donc préféré fixer dès le début la quantité d'absorption en suivant le modèle de Madau. J'ai testé l'effet de cette absorption en utilisant trois cas : l'absorption préconisée par Madau, puis moitié moins, et enfin deux fois plus que le facteur de Madau. Et pour chaque cas j'ai comparé le nombre de quasars correctement identifiés par HyperZ avec une estimation du décalage spectral relativement fiable (typiquement  $\Delta z < 0.8$ ). J'ai effectué les tests sur la base de quasars du DR3. Au final le nombre de quasars retrouvés avec une bonne estimation du décalage photométrique, pour des quasars à  $z > 2$  est meilleure en utilisant une plus petite quantité d'absorption que Madau : 6762 sur 9315 quasars retrouvés (parmi les 9768 quasars du DR3) dans ce cas contre 6162 sur 9211 quasars retrouvés avec plus d'absorption et 6769 sur 9291 quasars retrouvés avec moins d'absorption (j'ai privilégié l'exactitude

du décalage photométrique sur le nombre de candidats). A noter que Meiksin (2005) revoit l'absorption due au milieu intergalactique, mais la variation qu'il note par rapport à Madau est largement prise en compte par les tests que j'ai effectués, d'autant que la quantité d'absorption peut tout à fait varier d'une ligne de visée à une autre. Je reviendrai sur cette quantité d'absorption dans la partie sur le choix des templates.

### Pas du décalage spectral

L'un des paramètres importants dont a besoin HyperZ est le pas du décalage spectral qui va lui permettre de 'décaler' les templates fournis et calculer l'absorption dans les forêts Lyman- $\alpha$  et  $\beta$ . D'un côté un pas trop grand nous fera manquer certains décalages spectraux, d'un autre côté un pas trop petit peut amener à des dégénérescence lors du calcul du  $\chi^2$  et à un temps de calcul très long. J'ai déterminé le pas à utiliser de façon empirique en utilisant des simulations de quasars avec des erreurs très faibles. A partir des simulations réalisées avec des erreurs photométriques qui leur sont associées de l'ordre de 0.001 j'ai utilisé HyperZ pour retrouver le décalage spectral des quasars simulés avec des pas de 0.005, 0.01, 0.1, et 1, impliquant des temps de calcul (temps CPU) de 490., 235., 41. et 6. secondes respectivement.

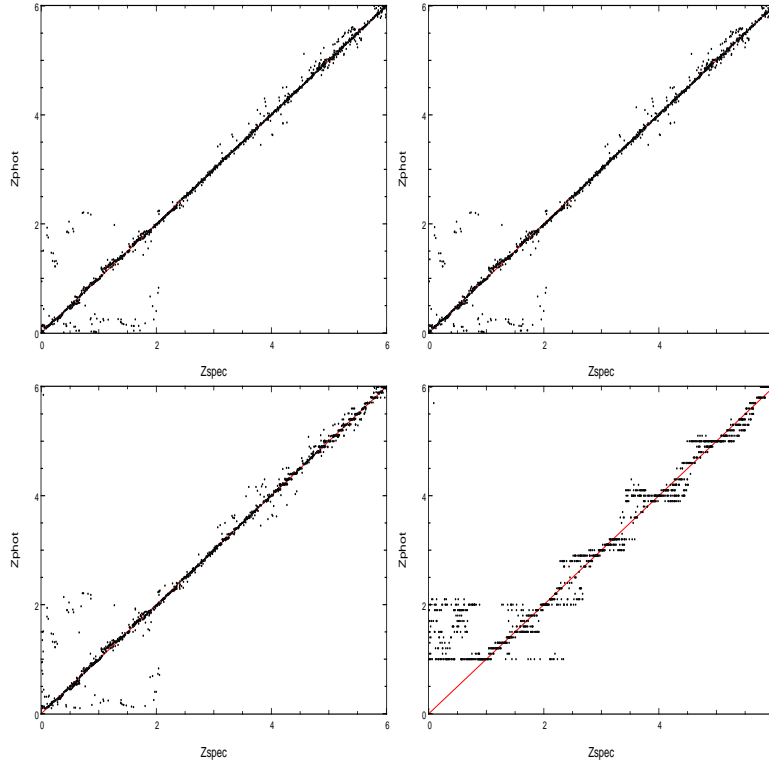


FIG. 4.7 – De gauche à droite et de haut en bas, décalage spectral photométrique en fonction du spectroscopique pour des pas de 0.005, 0.01, 0.1 et 1.

Au final, un pas de 0.01 a été retenu pour les quasars, car il représente le meilleur compromis entre temps de calcul et précision du décalage photomé-

trique calculé (cf. figure 4.7). Lors des tests avec des pas plus faibles, j'ai eu l'occasion de voir dans certains cas que le code de minimisation ne convergait pas. En effet, la recherche de  $\chi^2$  ne convergait pas, car les contraintes sur l'encadrement des valeurs de  $\chi^2$  étaient trop fortes, et il fallait au code, au bout d'un certain nombre d'itérations, recommencer la recherche. Afin d'éviter cela, même si j'utilise un pas plus grand, j'ai agrandi le champ de recherche de minima directement dans le code.

### Choix des templates

J'en arrive à une partie à la fois délicate et cruciale pour la sélection des objets : le choix des 'templates' c'est-à-dire du modèle qu'il va falloir comparer aux observations. De ce choix va dépendre les effets de sélection sur les candidats. J'entends par là, à la fois l'efficacité globale obtenue sur l'ensemble des candidats et aussi sur la nature des candidats (selon leur décalage spectral, selon la présence ou l'absence de certains systèmes d'absorption etc...) qui seront sélectionnés, ainsi que le type de contaminants retrouvés par l'analyse. Les templates se divisent en autant d'objets que l'on veut discerner, dans le cas de ma recherche il s'agit des templates d'étoiles et de quasars. En effet, comme je me limite aux objets ponctuels, il est inutile d'essayer de retrouver la population de galaxies qui se retrouverait dans ces objets ponctuels (leur nombre est négligeable). **La clé de la sélection est la mise en compétition de deux modèles : celui des étoiles et celui des quasars.** Néanmoins l'état de nos connaissances n'est pas le même pour chacun des deux. La population des étoiles étant bien mieux maîtrisée j'ai choisi dans un premier temps de prendre les templates d'étoiles utilisés par Hatziminaoglou et al. (2000), c'est-à-dire essentiellement les templates de Pickles (1998) et quelques spectres de naine blanche. Pour les quasars, les templates utilisés couvrent :

- Les templates de quasars utilisés par Hatziminaoglou et al. (2000)
- Une liste de templates réalisés par l'équipe d'HyperZ.
- Le spectre composite comprenant les données du SDSS (Vanden berk et al. 2001) et celle du HST (Telfer et al.2002) disponible à <http://www.pha.jhu.edu/rt19/composite/>
- Des templates issus du composite modifié au niveau des indices des lois de puissance (voir figure 4.8, les indices des pentes dans le bleu et le rouge de l'émission Lyman- $\alpha$  variant de : 0.3, 0.44, 0.75 et 0.9 puis 1.05, 1.1, 1.4, 1.76, 2.2 et 2.4 respectivement).

Dans les premiers tests j'ai mis en compétition ces différents modèles à ceux des étoiles en utilisant un catalogue issu de la base spectroscopique et un catalogue issu de la base photométrique.

Le taux de confirmation (TxC) d'une sélection se définit comme étant le nombre de quasars retrouvés sur le nombre de candidats. On peut définir ce taux pour un template :

$$TxC = \frac{VraiQSO}{NombreCandidats} \quad (4.11)$$



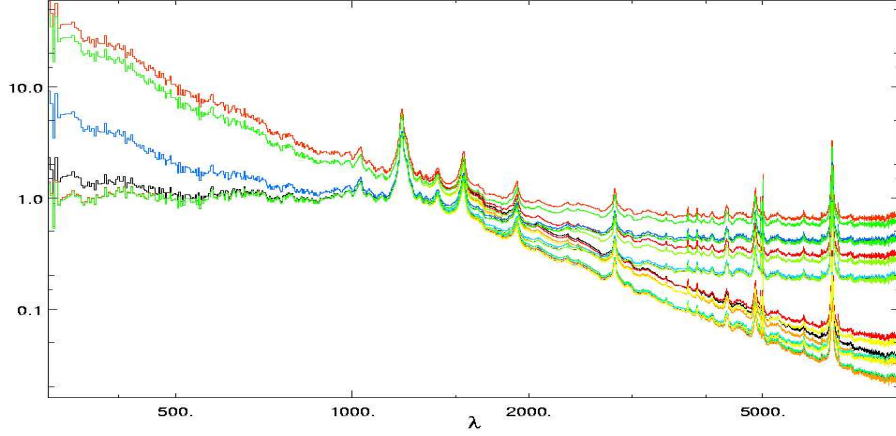


FIG. 4.8 – Template composite SDSS+HST avec les indices des pentes dans le bleu et le rouge de l'émission Lyman- $\alpha$  variant de : 0.3, 0.44, 0.75 et 0.9 puis 1.05, 1.1, 1.4, 1.76, 2.2 et 2.4 respectivement.

avec  $\text{NombreCandidats} = \text{MauvaisCandidats} + \text{VraiQSO}$  ; le nombre de mauvais candidats correspond au nombre de contaminants dans la sélection. L'inverse du taux de confirmation s'écrit alors :

$$\frac{1}{TxC} = 1 + \frac{\text{MauvaisCandidats}}{\text{VraiQSO}} = 1 + \text{templetude} \quad (4.12)$$

Donc si on requiert un taux de confirmation égal à 1, la templetude (cf. paragraphe suivant) devrait être nulle, cette équation étant valable pour chaque template.

J'ai utilisé l'ensemble de la base des templates à ma disposition, et j'ai regardé le nombre de quasars que chaque template permet alors de retrouver. Il y a donc aussi une confrontation entre les modèles de quasars.

Pour caractériser l'efficacité d'un template, j'ai défini le rapport entre le nombre de mauvais candidats et le nombre de vrais quasars et ce pour chaque template (appelé templetude). Plus un template sera efficace, plus sa templetude sera faible. Pour sélectionner quels templates seront utiles à l'identification, il faut mettre en compétition tous les modèles proposés.

En effet, lorsque les templates décrivent de trop près les différentes variations d'un même type d'objet (ici les quasars), et tentent de représenter autant de types de quasars que possible, on s'attend à voir une dérive entre classifier correctement les quasars un peu étranges et introduire plus de confusion pour les quasars 'normaux' (i.e. plus proches de la moyenne). Il convient donc de limiter l'échantillon de templates à une population 'moyenne', et l'objectif de ces tests est de trouver l'échantillon de templates qui recouvrent au mieux cette moyenne.

Le processus de sélection s'est déroulé par itération : je mets en entrée d'HyperZ un champ d'observations qui sera comparé à tous les templates disponibles comme décrit précédemment. Avec les résultats fournis par HyperZ (essentiellement les valeurs de  $\chi^2$  associées aux objets) je construis l'histogramme du

nombre de 'mauvais' candidats sur le nombre de vrais quasars retrouvés par les templates, une barre de l'histogramme correspondant à un template.

Il s'agit d'un histogramme normalisé, c'est-à-dire que d'une part le nombre de mauvais candidats est divisé par le nombre total de candidats. D'autre part le nombre de vrais quasars retrouvés par un template est divisé par le nombre de vrais quasars retrouvés par tous les templates, ce qui revient à multiplier l'équation 4.12 par le taux de confirmation de la sélection.

Ceci étant réalisé, j'enlève de la liste des templates utilisés par HyperZ le (ou les) templates les moins efficaces (ayant la complétude la plus grande), et je relance HyperZ avec cette nouvelle liste de templates. Il s'agit d'enlever les templates ayant un faible taux de confirmation, tout en gardant la complétude grande (qui se définit comme étant le nombre de quasars retrouvés sur le nombre de quasars attendus). En pratique je retire les plus mauvais templates jusqu'à obtenir une variation de 1% de la complétude par rapport à celle trouvée en utilisant tous les templates. Je réalise cette itération jusqu'à ce que le nombre de quasars retrouvés commence à diminuer, car cela signifie que certaines observations de quasars ne peuvent plus être reproduites par les modèles correspondant aux templates de quasars.

J'ai appliqué cette méthode de sélection sur la base spectroscopique du DR3 (selon les critères donnés en 2.1.2) : il s'agit de la région  $170 < ra < 185$  et  $-7.5 < dec < 7.5$  qui comprend 2867 objets spectroscopiques dont 1634 quasars (néanmoins l'intersection de l'ensemble "catalogue d'objets photométriques ayant un spectre dans cette région" et l'ensemble "quasars issus de la base spectroscopique dans cette région" ne comprend que 1366 objets ; ceci vient des critères de sélection sur les objets photométriques : cf. paragraphe 4.4 de Abazajian et al. (2004) décrivant le DR2. En résumé certains objets de la base spectroscopique ne sont pas présents dans la base photométrique). Je l'ai aussi appliquée sur un champ issu de la base photométrique : il s'agit de la région :  $183 < ra < 193$  et  $-4.5 < dec < 4.5$ , qui comprend 206557 objets dont 827 quasars répertoriés dans la base spectroscopique (l'intersection est de 692 quasars). La liste de template issue de la base spectroscopique du DR3 est clairement optimisée pour retrouver un maximum de quasars, tandis que celle issue de la base photométrique tente de minimiser le nombre de candidats (je rappelle que pour ce dernier cas, les objets sans identification spectrale ne sont pas considérés comme des quasars).

Afin de vérifier ces sélections il est nécessaire de tester les templates sur d'autres champs spectroscopiques et photométriques. J'ai utilisé un champ photométrique ( $i < 19.1$ ) contenant 559 quasars, pour les templates issus des tests de la base spectroscopique je retrouve 555 quasars pour 1843 candidats, tandis que pour la base photométrique je retrouve 549 quasars pour 1306 candidats. Ces vérifications confirment que les deux échantillons de templates sont optimisés l'un pour un maximum de quasars et l'autre pour minimiser le nombre de candidats tout en gardant un maximum de quasars parmi les candidats.

Ensuite j'ai voulu comprendre le rôle des templates d'étoiles, sur deux points, d'un coté quelles étoiles sont les principaux contaminants à la sélection de quasars et ensuite quelle est l'influence de l'absence ou de la présence de certains templates d'étoiles sur les mauvaises identifications, mais cette fois aussi bien

dans le cas de quasars identifiés comme étoiles que l'inverse.

Pour étudier le premier point j'ai travaillé avec une base composée de quasars uniquement, celle de Schneider et al. (2005) issue du DR3 du SDSS. J'ai tracé l'histogramme du nombre de quasars qui se voyaient attribuer un statut d'étoile, et ce en fonction du template d'étoile affecté à l'identification (cf. 4.9).

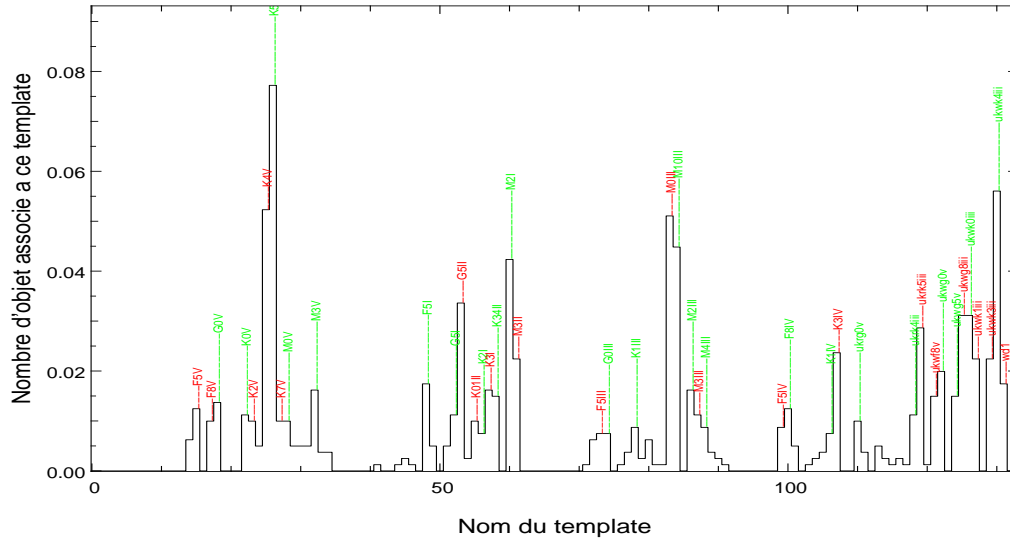


FIG. 4.9 – Histogramme normalisé des templates des étoiles associés aux quasars issus du DR3 qui ne sont pas reconnus comme tels par HyperZ.

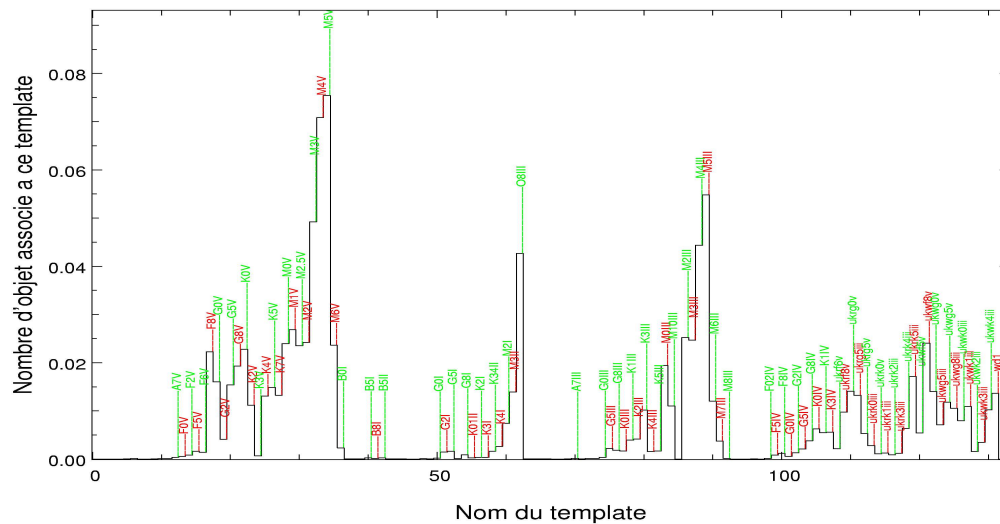


FIG. 4.10 – Histogramme normalisé des templates des étoiles associés correctement aux étoiles d'un champ du SDSS.

Ceci me donne une indication sur les contaminants que je vais retrouver, en effet les spectres retrouvés ici sont ceux qui sont les plus proches des spectres des quasars. Pour confirmer ceci et passer au second point j'ai effectué des tests sur une base spectroscopique du SDSS, qui contient aussi les spectres d'étoiles.

La figure 4.10 représente les templates d'étoiles qui permettent d'identifier correctement les étoiles (et donc de différencier les étoiles des quasars). Et il est visible en comparant les figures 4.9 et 4.10 que les templates d'étoiles qui participent à la mauvaise identification des quasars permettent aussi d'un autre côté d'identifier correctement des étoiles.

Ce faisant, la diversité des templates d'étoiles permet de limiter le nombre de candidats. En effet, on ne s'attendait pas à améliorer la sélection en enlevant des templates d'étoiles. La plupart des quasars étaient déjà retrouvés par la sélection simple, donc le nombre de quasars à récupérer en enlevant des modèles d'étoiles est très faible et coûtera très cher en nombre de contaminants supplémentaires. Les spectres d'étoiles ayant des spectres proches de ceux des quasars sont représentés figure 4.11 pour les types K5V et M0III, à côté du type d'étoile O8III qui lui est très éloigné du spectre des quasars.

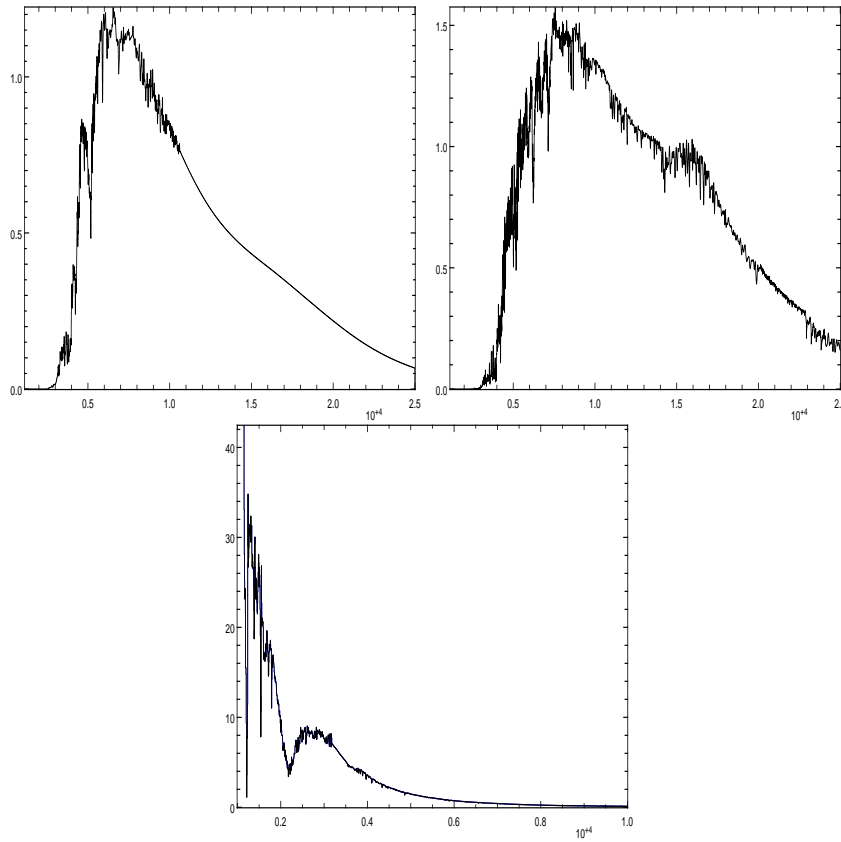


FIG. 4.11 – En haut, les types d'étoiles proches des quasars par leur couleur : K5V (la fin du spectre est interpolée) et M0III. En bas, spectre du type O8III dont la couleur est facilement différenciée de celle des quasars. Template de Pickles (1998)

Au final il faut vérifier au maximum l'hypothèse selon laquelle on connaît tous les spectres des objets à étudier, et cela est surtout vrai pour les étoiles. Les contaminants sont bien connus, mais il est essentiel de tous les avoir afin d'éviter les mauvais candidats quasars. Il faut donc couvrir au maximum la diversité des

spectres d'étoiles tout en évitant les redondances.

### Décalage spectral pour les tests

Une fois que les templates ont été choisis pour la sélection, il faut estimer la précision de la détermination des décalages spectraux photométriques. D'une part il s'agit de savoir s'il est possible d'utiliser cette information pour affiner la sélection des candidats, et d'autre part il faut chercher les dégénérescences qui peuvent être dues à un manque d'information dans les templates. Pour cela il est nécessaire de voir d'un côté quel décalage spectral est affecté aux 'bons' candidats quasars et avec quelle précision, et d'un autre côté quel est celui affecté aux 'mauvais' candidats. Pour la première partie j'ai de nouveau utilisé l'échantillon de quasars du DR3 et déterminé quel était le pourcentage de quasars correctement identifiés ayant une erreur sur leur décalage spectral inférieure à 0.5 (en utilisant HyperZ avec une quantité d'absorption inférieure à celle proposé par Madau (1995)). Ceci est représenté figure 4.12 sur laquelle est aussi représenté l'histogramme de la fraction de candidats ayant une différence entre décalage photométrique et spectroscopique inférieure à 0.5. Au décalage spectral  $z = 2.1$  le taux de quasars avec un bon 'décalage spectral' passe de 33% à plus de 70%. Les décalages photométriques estimés au-delà de 2.1 sont donc fiables à 0.5 près, c'est-à-dire lorsque l'émission Lyman- $\alpha$  rentre dans le domaine couvert par les filtres du SDSS.

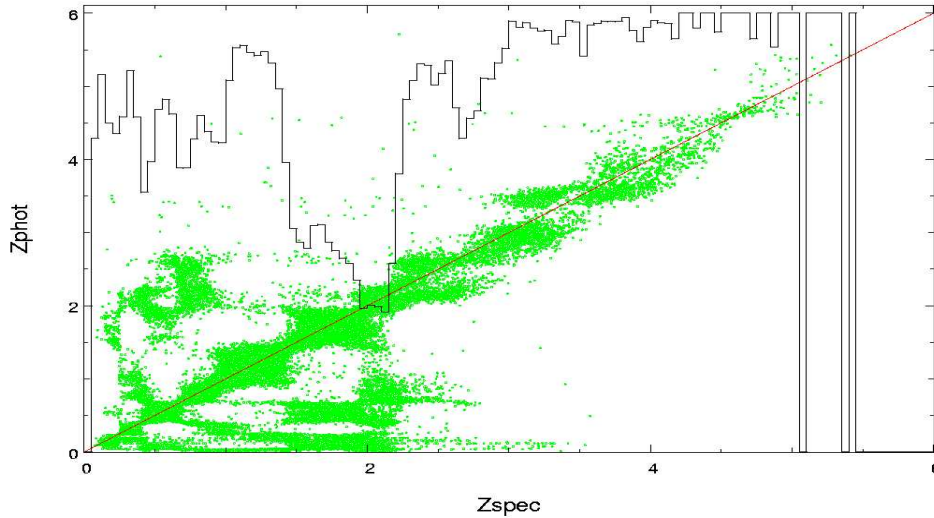


FIG. 4.12 – Décalage photométrique en fonction du décalage spectral spectroscopique à partir de l'échantillon de candidats quasars identifiés dans le catalogue de quasars du DR3 (donc les candidats sont des quasars). L'histogramme représente la fraction de candidats quasars ayant une différence entre décalage photométrique et spectroscopique inférieure à 0.5 (100% étant ramené à l'échelle de valeur 6.).

Les plus gros problèmes de dégénérescence se retrouvent pour des quasars ayant un décalage spectral entre 1.4 et 2.1 ; à partir de ce décalage spectral

l'émission Lyman- $\alpha$  du quasar rentre dans le domaine de longueur d'onde couvert par le filtre  $u$ . Il est en effet aisé d'attribuer un décalage spectral selon la position de cette raie dans chacun des filtres. L'erreur sur les décalages spectraux est ensuite fonction de la largeur de chacun des filtres et de la capacité du code à différencier cette raie d'autres raies du spectre.

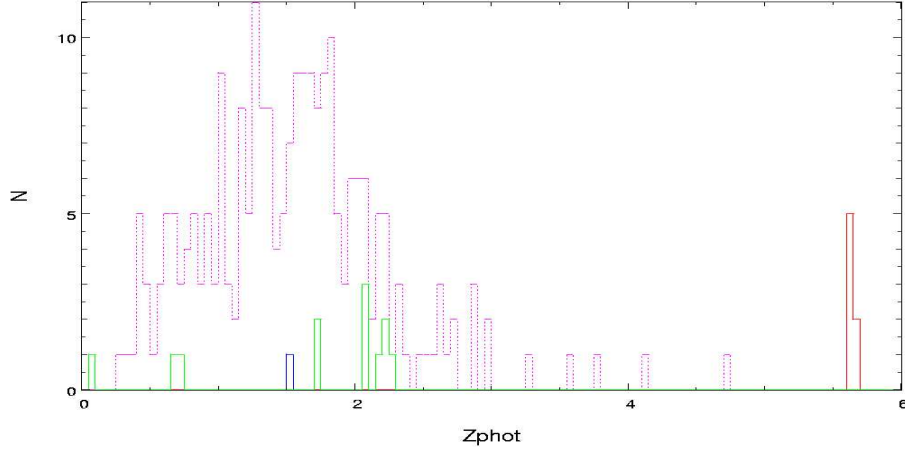


FIG. 4.13 – Histogramme des candidats quasars dans un champ du SDSS, selon leur nature : des quasars (magenta pointillé) ou des contaminants (en bleu les galaxies, en vert les étoiles, en rouge les étoiles plus évoluées). L'histogramme est en fonction du décalage spectral photométrique qui leur est associé par HyperZ.

Pour vérifier le classement des contaminants en fonction du décalage photométrique qui leur est attribué, j'ai tracé l'histogramme des candidats quasars dans un champ du SDSS ( $i < 19.1$ ), selon leur nature : des quasars ou des contaminants ; et ceci en fonction du décalage spectral photométrique qui leur est associé par HyperZ (cf. figure 4.13). Le champ comprend 182 quasars ; HyperZ retrouve 181 de ces quasars sur 441 candidats, mais seulement 195 ont une identification spectrale, ce qui signifie que 246 candidats sont de nature inconnue, la majorité correspond probablement à des étoiles. Néanmoins on peut tirer quelques conclusions sur l'échantillon spectral dont nous disposons : d'après la figure 4.13 les décalages spectraux attribués aux mauvais candidats quasars (en effectuant le test sur la base spectroscopique), se situent dans deux intervalles. Un premier entre 2 et 2.5 où les contaminants sont des étoiles normales et un autre intervalle avec un décalage au-delà de 5 qui comprend essentiellement des étoiles évoluées. Il est important de noter ici qu'il n'y a pas de quasar avec un décalage supérieur à 5. En effet, la probabilité d'avoir un contaminant plutôt qu'un quasar avec un tel décalage spectral est très importante, ce qui nous sera utile pour déterminer la sélection la plus efficace (voir dans le prochain paragraphe).

### Critères de sélection

Une fois qu'HyperZ a délivré les résultats, c'est-à-dire calculé un décalage spectral en lui associant une valeur de  $\chi^2$  et ce pour chaque modèle, il faut

pouvoir utiliser ces résultats pour déterminer si cela correspond à l'identification d'un quasar ou d'une étoile. La question à laquelle il faut répondre est : 'comment utiliser les résultats fournis par l'adaptation de DSE pour sélectionner les candidats ?'

Le choix le plus simple revient à comparer les valeurs des  $\chi^2$  calculées pour chaque modèle et pour un objet donné et estimer que le  $\chi^2$  le plus faible correspond au modèle le plus proche, et c'est ce qui a été fait jusqu'ici lorsque je parlais de candidats. Néanmoins à cause de l'existence de dégénérescences dans les valeurs des  $\chi^2$ , la valeur la plus faible ne correspond pas forcément au meilleur choix. Ces dégénérescences sont dues au fait que deux minima très proches peuvent au final amener à sélectionner le 'mauvais minimum'. C'est pourquoi certains auteurs (e.g. Weinstein et al. 2004) préfèrent estimer la probabilité associée à ce  $\chi^2$  et effectuer une discrimination à partir de cette probabilité.

Ils peuvent utiliser une telle option car leur traitement s'est toujours effectué avec des données complètement homogènes. HyperZ (cf. section 4.1), propose un traitement équivalent. En effet, la partie de minimisation de  $\chi^2$  permet de rechercher différents minima et de sélectionner le 'meilleur'. D'un autre côté HyperZ permet aussi de calculer les valeurs de probabilité associées aux  $\chi^2$  trouvés :

$$\begin{aligned} Q(\chi^2|\nu) &= 1 - P(0.5\nu, 0.5\chi^2) \\ &= \frac{\Gamma(0.5\nu, 0.5\chi^2)}{\Gamma(0.5\nu)} \\ &= \frac{1}{\Gamma(0.5\nu)} \int_{0.5\chi^2}^{\infty} e^{-t} t^{-1} dt \end{aligned} \quad (4.13)$$

avec  $\Gamma(k) = \int_0^{\infty} e^{-t} t^{k-1} dt$   $k > 0$  et avec  $\nu$  le nombre de degrés de liberté.

Néanmoins il faut garder à l'esprit que le but est de faire une sélection dans un autre relevé (le CFHTLS) que celui qui a permis la calibration. D'une part le passage de la photométrie du SDSS à celle du CFHT changera l'échelle de comparaison puisque les caractéristiques des deux relevés ne sont pas identiques. Et les erreurs photométriques dans ce dernier relevé n'ont pas fait l'objet de vérifications poussées (section 2.2 et Scranton et al. 2006 (<http://arxiv.org/abs/astro-ph/0508564> "Photometric Covariance in Multi-Band Surveys : Understanding the Photometric Error in the SDSS"), et comme dit le 'Numerical recipes in C' chapitre 15 page 660 :

This computed probability gives measure for the goodness-of-fit of the model. If  $Q$  is a very small probability for some particular data set, then the apparent discrepancies are unlikely to be chance fluctuations. **Much more probably either (i) the model is wrong-can be statistically rejected, or (ii) someone has lied to you about the size of the measurement errors  $\sigma_i$ -they are really larger than stated.**

ce qui signifie que d'une part les tests avec le SDSS permettent d'exclure la remise en cause des modèles utilisés et que d'autre part il est préférable si l'on ne veut pas accroître la dépendance de la sélection aux erreurs photométriques de réduire l'utilisation des probabilités. Dans cette optique j'ai préféré utiliser des

critères de sélection en me basant principalement sur les valeurs de  $\chi^2$ . Notez que ces valeurs dépendent bien sûr des erreurs. Je reviendrai sur leur influence et sur les solutions pour s'en débarrasser dans une certaine mesure.

Le choix de la sélection dépend de ce que l'on veut optimiser : soit le taux de confirmation (vrais quasars effectivement sélectionnés sur le nombre de candidats) soit la complétude (vrais quasars effectivement sélectionnés sur le nombre de quasars attendus). Dans les deux cas il faut néanmoins maximiser l'efficacité qui est définie comme étant le produit de la complétude et du taux de confirmation.

J'ai mis en place dans un premier temps trois types de sélection. Tout d'abord une sélection qui est spécialement optimisée pour le SDSS et permet d'avoir une très bonne efficacité :

(i) Une probabilité de rejet en tant qu'étoile certaine à 95%  
**et** pas de probabilité de rejet en tant que quasar certaine à 95%  
**et** une valeur de  $\chi^2$  de quasar inférieure à 60  
**et** un décalage spectroscopique associé au quasar inférieur à 5  
**ou**  
 une valeur de  $\chi^2$  associée au quasar inférieure à 1  
**et** une valeur de  $\chi^2$  associée aux étoiles supérieure à 3 avec un décalage spectral (déterminé par Hyperz) se situant entre 2.6 et 5.

La seconde partie des conditions (après le '**ou**') pour la sélection est essentielle pour pouvoir retrouver les quasars à grand décalage spectral. En effet, cet intervalle de décalage spectral contient des quasars auxquels sont associés des  $\chi^2$  très faibles, mais il y a aussi des étoiles avec un  $\chi^2$  relativement petit, c'est pourquoi j'impose une limite supérieure sur la valeur de  $\chi^2$  pour les quasars et inférieure pour les étoiles dans un certain intervalle de décalage spectral.

Une autre sélection adaptée dans une certaine mesure au SDSS correspond à la même sélection présentée mais en y enlevant les références directes aux valeurs des  $\chi^2$  :

(ii) Une probabilité de rejet en tant qu'étoile certaine à 95%  
**et** pas de probabilité de rejet en tant que quasar certaine à 95%  
**et** un décalage spectroscopique associé au quasar inférieur à 5  
**ou** une valeur de  $\chi^2$  associée aux quasars inférieure à celle associée aux étoiles.

Ceci diminue l'efficacité, en augmentant le nombre de candidats mais reste comparable à l'optimisation SDSS, la différence primordiale se situant au niveau des quasars à plus grand décalage spectral subissant une plus forte contamination.

Enfin une sélection basique :

(iii) Une valeur de  $\chi^2$  associée aux quasars inférieure à celle associée aux étoiles  
**et** avec un décalage spectroscopique inférieur à 5.

Néanmoins pour toutes ces sélections, on retrouve un problème dû au fait que les couleurs de certains quasars sont très proches du locus des étoiles. Il est possible de rendre compte de ce phénomène en traçant l'histogramme des candidats quasars en fonction du décalage photométrique déterminé par HyperZ,



cf. figure 4.14. Notez que cet effet est surtout visible dans le cas d'un test de sélection proche des conditions réelles, c'est-à-dire sur la base photométrique qui contient une proportion réelle d'étoiles. La figure montre un pic de candidat avec un décalage photométrique compris entre 1.9 et 2.2, ce qui correspond aux limites entre couleurs de quasars et locus des étoiles, zone où il est difficile de distinguer les deux types d'objet. Afin de tenir compte des difficultés que

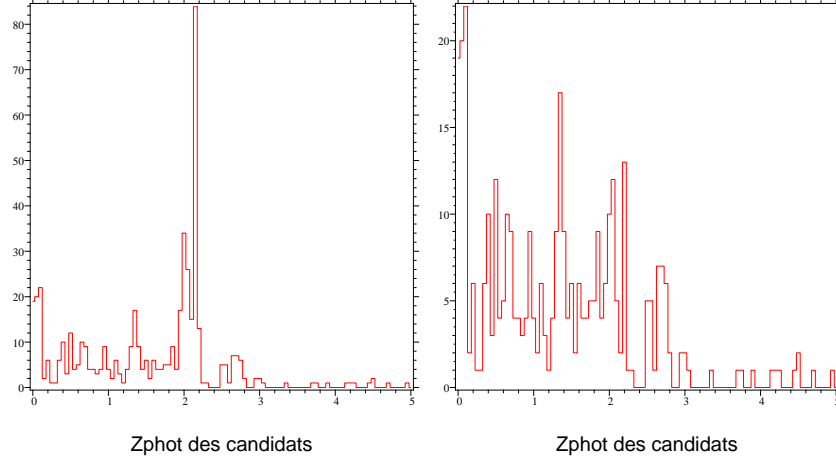


FIG. 4.14 – Histogramme des candidats quasars en fonction du décalage spectral photométrique qui leur est associé par HyperZ. Sélection réalisée dans la base photométrique. Panneau de gauche : sélection basique. Panneau de droite : contraintes sur le rapport  $\chi^2_{qso}/\chi^2_{etoile}$  : inférieur à 0.3.

rencontre HyperZ pour sélectionner correctement les quasars dans cette zone, j'ai choisi de contraindre plus fortement le rapport  $\chi^2_{qso}/\chi^2_{etoile}$ . Après différents tests, j'estime que celui-ci doit être de 0.3, afin de limiter le nombre de candidats dans cette zone mais sans pour autant manquer les vrais quasars. Cette sélection peut se résumer ainsi (sélection avancée) :

**Si**  $z_{phot}$  compris entre 1.9 et 2.2 :

Une valeur de  $\chi^2$  associée aux quasars inférieure à celle associée aux étoiles

**et**  $\chi^2_{qso}/\chi^2_{etoile} < 0.3$

**Si non** :

Une valeur de  $\chi^2$  associée aux quasars inférieure à celle associée aux étoiles

**et** avec un décalage spectroscopique inférieur à 5.

La table 4.3 récapitule l'efficacité, la complétude et le taux de confirmation pour chacune de ces sélections sur la base de données de quasars de Schneider et sur la base photométrique du SDSS. Ces statistiques ont été réalisées avec pour modèles l'échantillon de templates sélectionnés via la base spectroscopique et une absorption inférieure à celle de Madau (1995). Les objets sélectionnés ont une magnitude limite de 19.1 afin d'avoir une estimation de la complétude qui peut être réalisée à partir du nombre de quasars effectivement détectés dans le SDSS, pas par rapport au nombre de quasars attendus. La base photométrique fait 24 degrés carrés, et a été choisie pour avoir la plus grande partie de ses objets

Base utilisée	type de sélection	taux de confirmation	Complétude	efficacité
Base de quasar du DR3				
	Sélection optimisée SDSS (i)	40669/40669 <b>100 %</b>	40669/46420 <b>87.6 %</b>	<b>87.6 %</b>
	Sélection large SDSS (ii)	45612/45612 <b>100 %</b>	45612/46420 <b>98.3 %</b>	<b>98.3 %</b>
	Sélection basique (iii)	45609/45609 <b>100 %</b>	45609/46420 <b>98.3 %</b>	<b>98.3 %</b>
	Sélection avancée	45520/45520 <b>100 %</b>	45520/46420 <b>98.0 %</b>	<b>98.0 %</b>
Base photométrique				
	Sélection optimisée SDSS (i)	173/249 <b>69.5 %</b>	173/182 <b>95.1 %</b>	<b>66.1 %</b>
	Sélection large SDSS (ii)	181/441 <b>41.0 %</b>	181/182 <b>99.5 %</b>	<b>40.8 %</b>
	Sélection basique (iii)	181/435 <b>41.6 %</b>	181/182 <b>99.5 %</b>	<b>41.4 %</b>
	Sélection avancée	181/316 <b>57.3 %</b>	181/182 <b>98.5 %</b>	<b>56.4 %</b>

TAB. 4.3 – Évaluation du taux de confirmation, de la complétude et de l’efficacité (voir texte), pour quatre types de sélection différentes appliquées sur deux bases, l’une uniquement de quasars, l’autre photométrique.

suivis en spectroscopie. Le catalogue photométrique contient  $\sim 35500$  objets à la magnitude limite  $i < 19.1$ . Au final, lors du passage d’un catalogue de quasars à la base photométrique je retrouve toujours une grande partie (95%) des quasars mais avec deux fois plus de candidats (taux de confirmation de 56.4% en sélection avancée), ces candidats ayant pour la plupart (75%) une nature inconnue.

Pour l’instant il ne s’agit pas de définir immédiatement quelle sélection sera retenue au final mais de comprendre l’influence de chacune sur le nombre et la nature des candidats. Les tests sur la base photométrique sont ceux qui sont les plus proches de la sélection finale qui sera réalisée ; les conditions sont les plus proches des difficultés auxquelles je serai confronté.

Dans cette sélection (la (i)), parmi les 78 faux candidats, seulement 16 d’entre eux ont un spectre, leur nature est donc inconnue, j’ai supposé qu’ils n’étaient pas des quasars. Leur position dans les diagrammes couleur est représentée dans la figure 4.15. La plupart de ces candidats sont très proches du locus habituel des quasars, et certains d’entre eux ont des couleurs assez inhabituelles, en-dehors du locus des étoiles, que l’on voit en magenta sur la figure. Ainsi une petite partie de ces objets aux couleurs inhabituelles pourraient être rejetés si l’on couplait la sélection via HyperZ à un découpage dans les diagrammes couleur-couleur ; ce qui n’est pas forcément voulu si l’on veut aussi pouvoir trouver des objets rares avec des couleurs peu fréquentes. Notez que tous les objets issus de la base spectroscopique ont fait l’objet d’une sélection par couleur, notamment les quasars (décrit en détail en 3.3.4), mais aussi les étoiles.

A partir de ce moment lorsque je parlerai de sélection, il s’agira de la sélection optimisée pour les données issues du SDSS (la première présentée dans le tableau 4.3).

### Nature des candidats

La limite en magnitude de sélection des quasars pour l’algorithme de recherche de quasars du SDSS de Richards et al (2002) est de 19.1 pour la sélection normale et de 20.2 pour les quasars à grand décalage spectral (cf. 2.1). Mais pour réussir à atteindre une densité de quasars suffisante pour trouver des groupes suffisamment proches de quasars il faut aller à une profondeur en magnitude plus grande. J’ai donc voulu tester ma sélection au-delà des limites

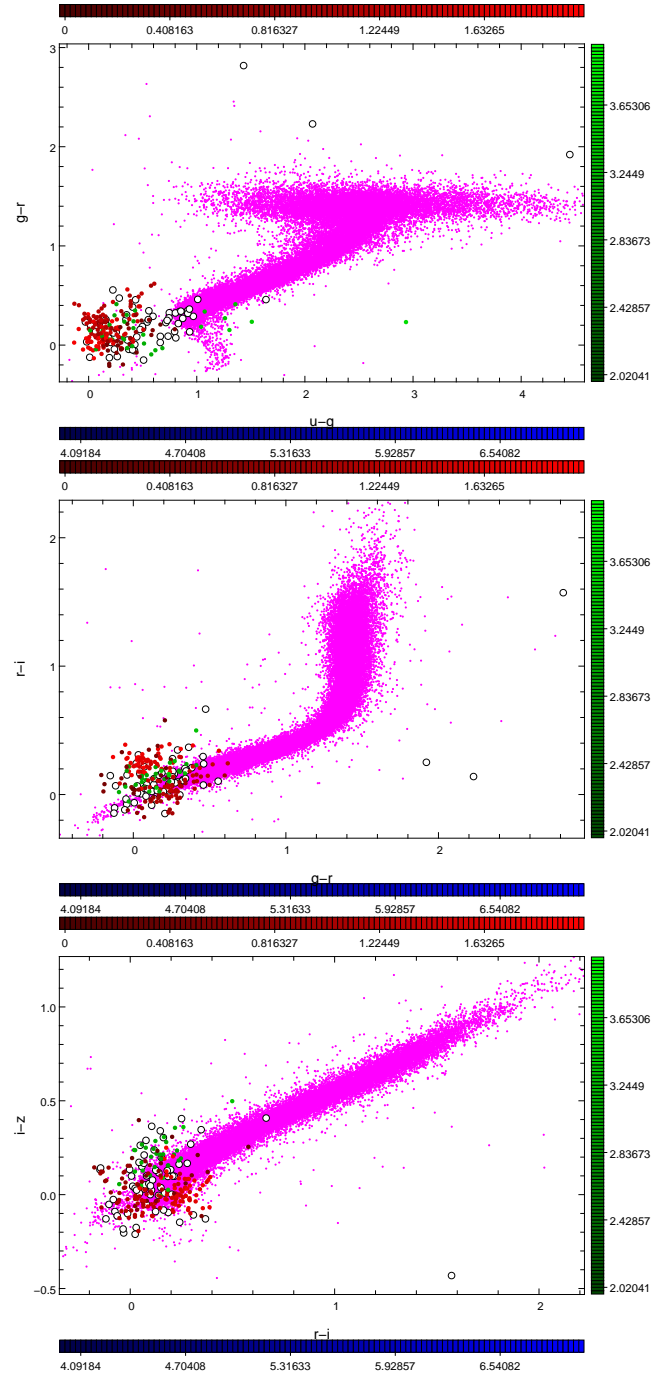


FIG. 4.15 – Diagramme couleur couleur, représentant en magenta l'ensemble des objets de la base photométrique, en couleur les quasars selon leur décalage spectral (voir échelle de couleur sur les cotés : en rouge  $z < 2$ , en vert  $2 < z < 4$ , en bleu  $z > 4$ ) et les candidats qui ne sont pas des quasars (cercles noirs). Les points en magenta représentent l'ensemble des objets de l'échantillon et on discerne donc essentiellement les étoiles.

imposées pour la sélection du SDSS, ceci afin d'avoir une idée de la densité de candidats obtenue, tout en gardant à l'esprit que la photométrie du SDSS ne permet pas d'aller au-delà de  $i=24$  en théorie.

En vérifiant le statut de ces nouveaux candidats dans la base de données du SDSS je me suis aperçu qu'une partie d'entre eux ( $\sim 30\%$ ) ont un drapeau *qso\_faint* ce qui signifie qu'ils ont été sélectionnés par le SDSS mais qu'ils ne répondent pas aux critères de magnitude, et qu'ils feront probablement l'objet d'un prochain suivi spectroscopique. En poussant ainsi jusqu'aux limites du relevé, les informations disponibles sur les objets sont trop peu nombreuses pour permettre une classification des objets, puisqu'aucun suivi spectroscopique n'a été réalisé.

Enfin la question que l'on se pose naturellement est : quels sont les quasars qui ne sont pas identifiés comme tels et pourquoi ? Pour répondre à cette question le mieux est de considérer la base spectroscopique et d'étudier en détail les spectres des quasars qui ne sont pas retenus comme candidats.

Il apparaît qu'une partie des quasars (20%) ratés sont de Type BAL, c'est-à-dire présentant de très fortes absorptions (Broad Absorption Line) à plusieurs milliers de kilomètres par seconde des raies d'émission, ce qui peut modifier leurs couleurs si ces absorptions sont suffisamment larges et se situent dans le filtre en même temps que l'émission. En effet, cela peut revenir à regarder un spectre plat (absorption + émission) à travers un filtre, en enlevant une partie du flux à travers ce filtre ; notamment dans la forêt Lyman- $\alpha$  où le flux moyen est plus faible que dans des spectres sans BAL.

Il en est de même pour les systèmes lorentziens : leur présence augmente la quantité moyenne d'absorption dans la forêt Lyman- $\alpha$ . Cela revient à surestimer le décalage spectral photométrique. Il est possible d'observer ce décalage des couleurs dans la figure 4.16. En noir est représentée la population 'normale' de quasars, avec un décalage spectral  $2 < z < 4$ , en bleu la population de quasars avec un système lorentzien issu du DR3 du SDSS et de Prochaska et al. (2005). Le diagramme  $g - r / u - g$  rend particulièrement bien compte de cet effet de décalage des couleurs. Ceci ne joue a priori pas sur l'identification à proprement parler des quasars, mais plus sur la détermination de son décalage spectral.

La plupart des quasars ratés ne peuvent pas être 'récupérés' c'est-à-dire retrouvés avec une sélection un peu plus large ou en effectuant une sélection précise sur leurs couleurs, car dans le plan des  $\chi^2$  ils se trouvent trop proches des étoiles. En effet, même dans les diagrammes couleur-couleur, on peut voir que la plupart de ces quasars qui ne sont pas sélectionnés sont proches du locus des étoiles, cf. figure 4.17. La figure représente en noir les quasars du DR3, et en couleur les quasars de cet échantillon qui ne sont pas reconnus comme tels par HyperZ (12% de l'échantillon). La plupart d'entre eux dessinent la limite entre le locus des quasars et celui des étoiles, ou bien croisent le locus des étoiles. Certains d'entre eux, néanmoins, sont assez loin de la bordure mais ne sont pas reconnus, c'est le prix à payer pour l'équilibrage entre taux de confirmation et complétude.

Néanmoins une méthode complémentaire pourrait permettre de les retrouver, notamment les BAL, qui ont la caractéristique de présenter de fortes variations. Donc une recherche via la variabilité en complément des couleurs permet-

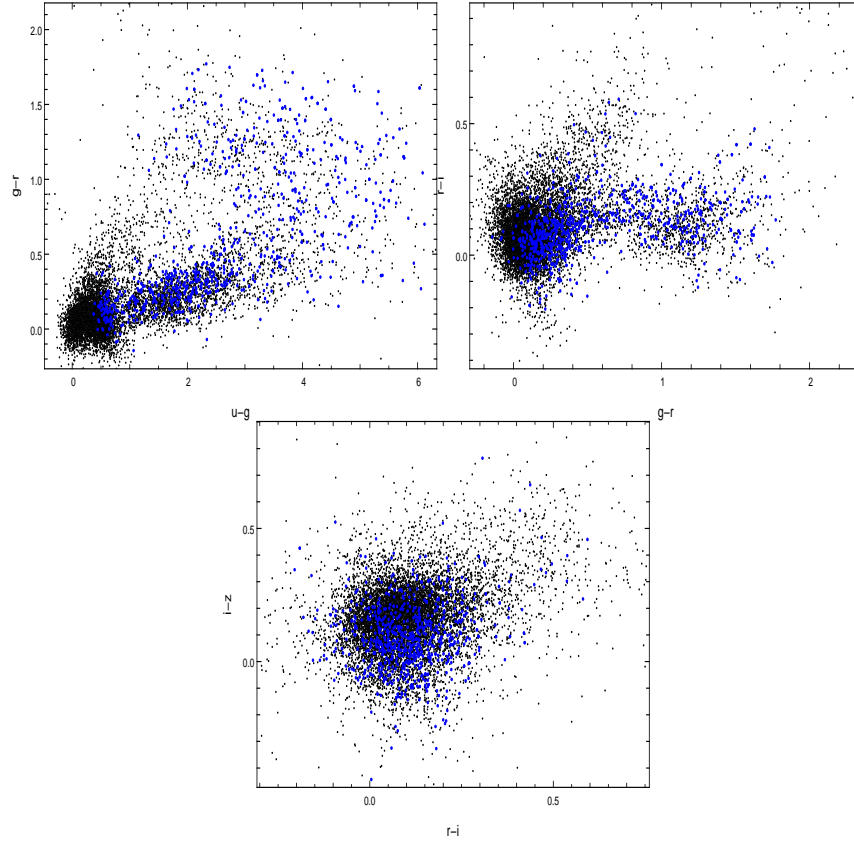


FIG. 4.16 – Diagramme couleur-couleur, représentant en noir l'ensemble des quasars du DR3 avec un décalage spectral supérieur à 2, en bleu les quasars présentant un système lorentzien dans leur spectre.

trait a priori d'en retrouver une partie, pour peu que l'on désire les récupérer. D'autre part la sélection du SDSS est enrichie du relevé FIRST, une partie des candidats SDSS sont ciblés grâce à ce relevé. La sélection proposée ici ne tient pas compte de cette recherche complémentaire (et si celle-ci devait s'enrichir d'information dans un autre domaine d'énergie, l'idéal serait le domaine X, non accessible avec les observatoires terrestres). La comparaison des deux sélections n'est donc pas si directe ; je ne propose qu'une sélection sur les données optiques tandis que le SDSS se base sur des données optiques et radio...

Maintenant que la méthode de sélection est définie il va falloir mesurer à quel point celle-ci est efficace pour le relevé du CFHTLS. En tenant compte de la différence des filtres, de la limite en magnitude, du type de photométrie utilisée et aussi du matériel que fournit le relevé, notamment les erreurs photométriques et les classements de qualité photométrique.

### 4.2.3 Sélection sur le CFHTLS

Les sélections que je vais décrire ici font référence au relevé du CFHTLS décrit dans la partie 2.2, et plus particulièrement aux champs profonds du relevé.

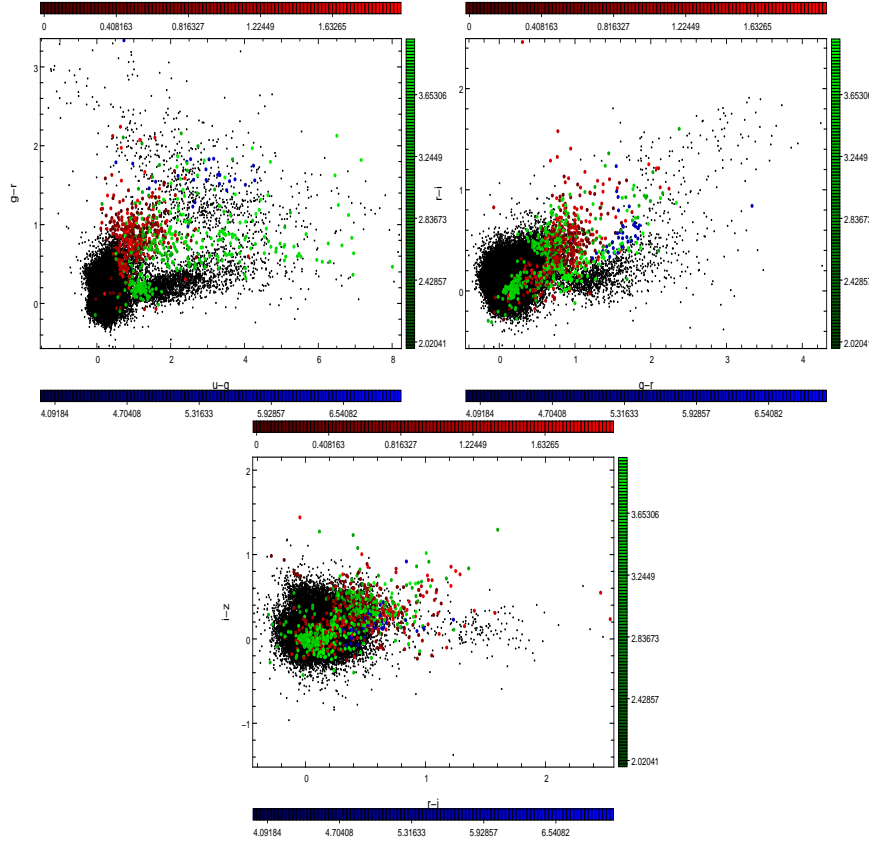


FIG. 4.17 – Diagramme couleur couleur, représentant en noir l'ensemble des quasars du DR3, en couleur les quasars non reconnus comme tels par HyperZ.

### Catalogues photométriques

La prise en main des données issues du CFHTLS s'est faite de manière progressive, au fur et à mesure que les données de chaque champ étaient complétées. L'historique des mise-à-jour des données ainsi que leurs différentes caractéristiques se trouvent dans la partie 2.2.

Tout d'abord quelles données issues des catalogues ai-je utilisées ? Comme déjà précisé précédemment, SourceExtractor produit des catalogues photométriques pour chaque filtre, notamment en utilisant l'astrométrie déterminée via une image trois-couleurs ( $g r i$ ). Ainsi les objets présents dans l'ensemble des catalogues doivent être détectés dans un de ces trois filtres. Les objets détectés uniquement en  $u$  ou  $z$  ne sont pas pris en compte. Il m'a fallu déterminer quels critères appliquer dans ces catalogues pour sélectionner les objets qui feront l'objet de la sélection via HyperZ : L'équipe de Térapix fournit un masque pour chaque champ de Megacam qui permet d'exclure les objets qui n'ont pas une photométrie correcte, par exemple les objets trop proches d'étoiles brillantes ou bien trop proches des bords de l'image. J'ai donc commencé par enlever les objets qui sont dans ces masques.

Le point le plus important est probablement la séparation entre galaxies et objets ponctuels. En effet, SourceExtractor permet d'attribuer une valeur à un

objet selon qu'il est plus ou moins proche, morphologiquement, d'une galaxie ou d'une étoile. Ce paramètre appelé 'CLASS\_STAR', varie entre '1' pour les étoiles et '0' pour les galaxies. Ce paramètre peut changer pour un catalogue issu d'un filtre à l'autre, mais il est le plus fiable dans le filtre  $i$ , j'ai donc choisi de prendre les objets dont la valeur dans ce filtre est supérieure à 0.5. Cela permet de sélectionner quelques quasars dont l'hôte pourrait être quelque peu résolu.

Ensuite SourceExtractor fournit un 'drapeau' sur la qualité photométrique de chaque objet extrait. Une valeur de '0' signifie qu'il n'y a aucun défaut apparent ; une valeur de 1 signifie que l'objet a probablement des voisins proches et brillants qui peuvent affecter la photométrie de l'objet (plus de 10% de la surface intégrée est concernée) ; une valeur de 2 signifie que l'objet était à l'origine fondu dans un autre. Les autres valeurs s'étendant jusqu'à 128 peuvent être trouvées dans le manuel (E. Bertin). Je m'arrête à cette description car ce sont les seuls drapeaux qui se retrouveront dans mes catalogues. En fait, à l'origine, j'avais opté pour un drapeau de valeur '0', mais vu le nombre de quasars connus concernés par les drapeaux de valeur '1' et '2', j'ai finalement accepté les drapeaux de valeur jusqu'à '3' (c'est-à-dire cumulant les défauts '1' et '2'), cette valeur devant être vérifiée pour les catalogues extraits dans chacun des filtres. En plus de ceci, je sélectionne les objets qui ont au moins dans un des filtres une erreur photométrique inférieure à 0.2. Celles-ci sont d'ailleurs modifiées pour tenir compte des erreurs de la calibration qui sont estimées à 0.05, cette valeur est donc ajoutée en quadrature aux erreurs fournies par SourceExtractor. D'après une étude réalisée par C. S. Stalin (communication privée), les objets ayant une magnitude en  $i$  inférieure à 17 sont probablement saturés. Cette limite a été estimée après avoir étudié le rayon des objets en fonction de leur magnitude dans chaque filtre, par exemple le filtre  $g$  en figure 4.18, où l'on remarque une augmentation soudaine du rayon pour des magnitudes en  $g = 16.5$ .

Une fois les objets extraits du catalogue il me faut calculer le dérougissement pour chacun d'entre eux. Je réalise ceci en utilisant la carte de Schlegel et al. (2003), qui fournit les valeurs de  $E(B-V)$  en fonction de la position. Néanmoins pour parvenir à la valeur de  $A_\lambda$  correspondant au filtre utilisé il est nécessaire d'utiliser une constante de conversion. Celle-ci n'a pas été calculée pour les filtres de Megacam, j'ai donc utilisé les valeurs issues des filtres du SDSS.

## Sélection

La sélection adoptée est la sélection dite 'avancée', décrite dans la section 4.2.2 et présentée en dernière dans la table 4.3 à ceci près que les conditions sur le décalage spectral photométrique ne sont pas les mêmes pour les couleurs issues du SDSS et celles issues du CFHTLS. En effet, comme on peut le voir sur la figure 4.19, essentiellement pour la couleur  $u - g$ , pour un même décalage spectral photométrique, la couleur  $u - g$  issue du SDSS et celle issue du CFHT sont légèrement décalées. La sélection se retrouve donc modifiée afin que le domaine de décalage spectral photométrique concerné se situe entre 2.2 et 2.4. Il se trouve *in fine* que la méthode de sélection à appliquer dépend du suivi spectroscopique. Mais au moment de l'écriture de ces lignes, notre première

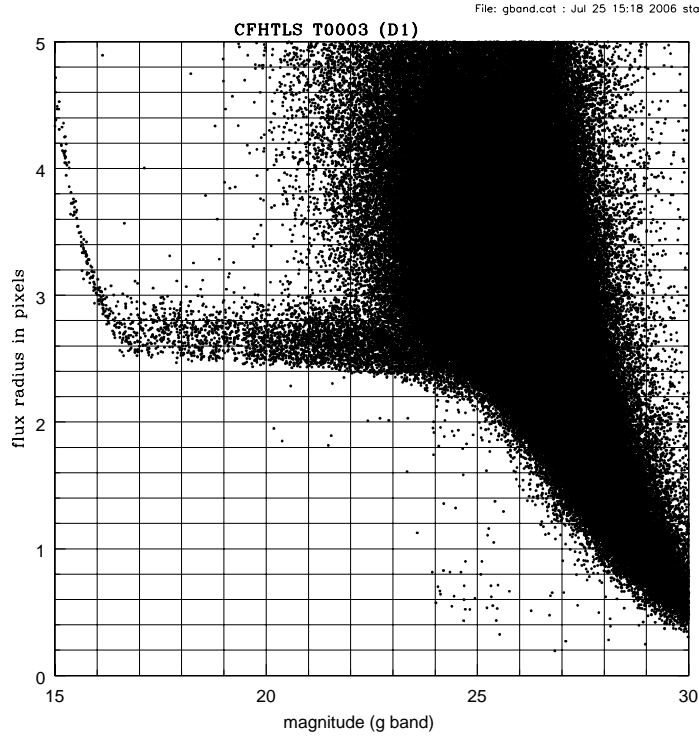


FIG. 4.18 – Évolution du rayon des objets en fonction de leur magnitude dans le filtre  $g$ . Figure fournie par C. S. Stalin.

stratégie de suivi spectroscopique n'a pas pu être appliquée et nous sommes en cours de réflexion pour un nouveau suivi (notamment en couplant différentes méthodes d'identification). La table 4.4 récapitule les résultats des candidats sur les différents champs du 'Deep'. Dans la plupart des champs, sauf le 'D1', un certain nombre de quasars existaient déjà dans la littérature (trouvés par exemple via [simbad.u-strasbg.fr](http://simbad.u-strasbg.fr)). En particulier pour les champ D2 et D3 suivis tous deux par le SDSS, et le champ D2 cible du 'cosmos field' (Prescott et al. 2006). Pour le D1 une liste de quasars a été publiée très récemment par l'équipe du VVDS (Gavignaud et al. 2006).

### Suivi spectroscopique

Une demande avait été faite sur les télescopes de l'ESO sur l'instrument VIMOS (VIsible MultiObject Spectrograph) monté sur Melipal (UT3) en utilisant sa fonctionnalité en MOS (Multi-Object Spectroscopy). Ce mode permet d'obtenir dans un champ donné un grand nombre de spectres d'objets et donc de les identifier pour un temps d'observation modique. Cette demande portait sur un des champs profonds du CFHTLS. Celle-ci a été refusée, arguant que des données en infrarouge ainsi qu'un couplage avec des sources en X permettrait d'accroître notre efficacité. Il nous faut donc attendre d'avoir ces données en infrarouge et de faire la corrélation avec des catalogues X pour déposer une nouvelle demande de temps. Un projet de suivi sur un seul champ avec VI-



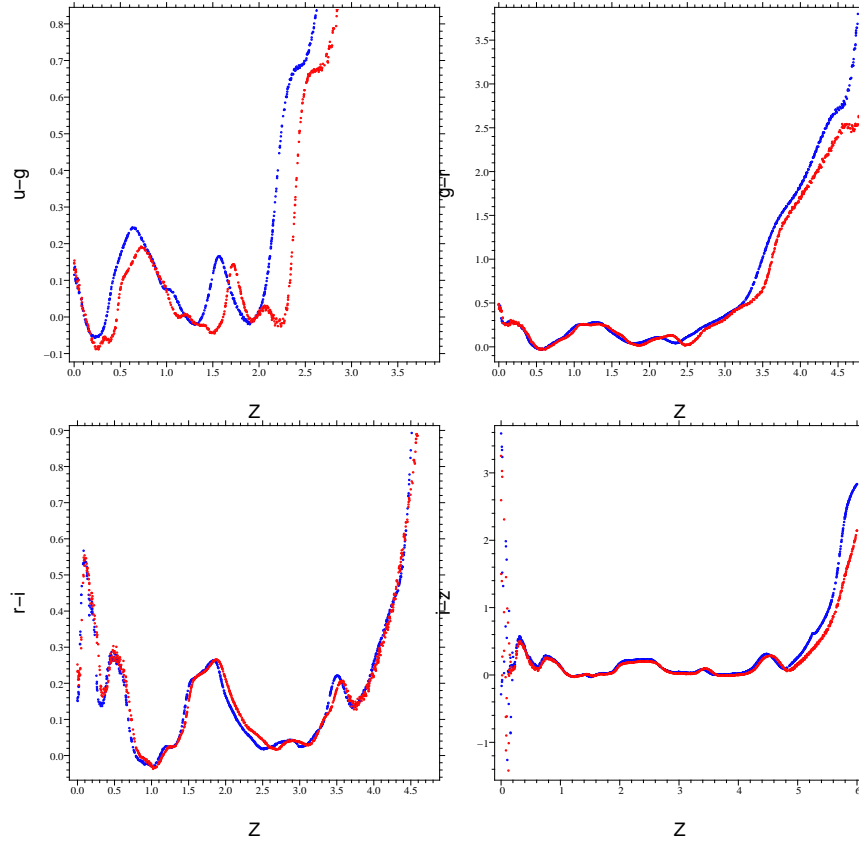


FIG. 4.19 – Évolution des couleurs des quasars (simulés) en fonction du décalage spectral. En rouge avec les couleurs des filtres du CFHT et en bleu celles issues des filtres du SDSS. Remarquez particulièrement le décalage pour la couleur  $u - g$ .

MOS nécessiterait une vingtaine d'heures, il est possible de réaliser un projet plus petit pour se concentrer sur des zones de regroupement de candidats, ou au contraire un projet plus grand en utilisant les capacités de haute résolution de l'instrument. Dans tous les cas la stratégie d'observation va dans un sens conditionner le type de sélection que nous aurons à utiliser.

Il existe un projet de suivi sur l'AAT (Anglo-Australian Telescope) mettant à contribution des sources en X pour une partie des champs W1 et D1 ainsi que D4. A terme le projet est d'aller jusqu'à une profondeur de  $g = 23$ . Mais un suivi spectroscopique jusqu'à ces magnitudes nécessitera probablement l'ELT (Extremely Large Telescope).

#### 4.2.4 Conclusion

Les tests menés sur les bases du SDSS nous permettent de disposer de plusieurs types de sélection selon les stratégies des suivis spectroscopiques. Pour chacune d'entre elles nous avons une estimation de l'efficacité. J'ai présenté comment ces critères de sélection pouvaient être appliqués au relevé CFHTLS

Champ cible	magnitude limite en g	nombre de candidats / nombre d'objets	Quasars identifié / quasar dans le champ	surface du champ (degré carré)
D1	19	47/367		
	20	69/635		
	21	118/1234		
	22.5	711/3710	27/31	0.77
D2	19	136/700		
	20	149/1125		
	21	215/1861		
	22.5	692/4571	70/72	0.69
D3	19	70/462		
	20	90/798		
	21	151/1458		
	22.5	693/3865	14/14	0.83
D4	19	211/926		
	20	231/1556		
	21	292/2654		
	22.5	847/6211	5/6	0.82

TAB. 4.4 – Récapitulatif des candidats pour les champs profonds du CFHTLS, avec une sélection de type 'avancée'. Pour certains champs il existe des catalogues de quasars, ceux-ci sont en grande partie retrouvés par ma sélection.

et quelles en étaient les contraintes. Sur certains champs profonds du CFHTLS nous disposons de petites bases de données de quasars qui nous permettent de quantifier dans une certaine mesure l'efficacité de notre méthode. Notamment en utilisant dans le champ D1 les relevés de XMM-Newton nous retrouvons près de 70% de ces sources parmi nos candidats (ce qui est un résultat attendu car 70% des sources détectées en X sont des noyaux actifs de galaxies). Au final la complétude de notre sélection est estimée à plus de 90% sur le relevé CFHTLS.



## Chapitre 5

# Milieu intergalactique

*"To make a bad day worse, spend it wishing for the impossible."*  
– Calvin

## Sommaire

---

<b>5.1 Observations</b>	<b>134</b>
5.1.1 Forêt Lyman- $\alpha$ en observation	135
5.1.2 Regroupement des raies	136
<b>5.2 Un milieu photo-ionisé</b>	<b>136</b>
<b>5.3 Modèle actuel</b>	<b>137</b>
5.3.1 La forêt Lyman- $\alpha$	138
5.3.2 L'enrichissement en métaux du milieu intergalactique	140
<b>5.4 Tomographie du milieu</b>	<b>140</b>
5.4.1 Utilisation de paires de quasars	141
5.4.2 Futur de la tomographie - les groupes de quasars	141
<b>5.5 Transverse correlation in the Lyman-<math>\alpha</math> forest</b>	<b>145</b>
5.5.1 Introduction	146
5.5.2 Observations	147
5.5.3 Numerical simulations	150
5.5.4 The observed correlation functions	151
Calculating correlation functions	151
The observed longitudinal correlation function	152
The observed transverse correlation function	152
Estimation of errors	153
5.5.5 Comparison of observations and simulations	154
Systematic uncertainties	154
Observed <i>vs</i> simulated correlation functions	156
5.5.6 Metal absorption systems	158
Identifying metal lines	158
The correlation of C IV systems	160
Peculiarities	163
5.5.7 Conclusions	164

---

## 5.1 Observations

Pour étudier le milieu intergalactique, qui ne peut pas être observé, nous avons vu dans le chapitre précédent qu'il est possible de le détecter indirectement en voyant de quelle manière il influence la lumière provenant de sources lointaines : les quasars. En effet, comme nous l'avons vu, le milieu intergalactique laisse son empreinte dans le spectre d'objets lointains, et cette empreinte contient des informations sur sa température et sa composition. A partir d'une ligne de visée en direction d'un quasar à grand décalage spectral il est possible d'observer la forêt Lyman- $\alpha$  (décrite dans l'introduction) qui s'étend sur de larges intervalles en décalage spectral. Ceci permet d'explorer une fraction très importante du temps de Hubble pendant lequel on s'attend à voir une évolution générale des absorbants, par exemple sur le nombre de raies d'absorption ou sur la profondeur optique moyenne.

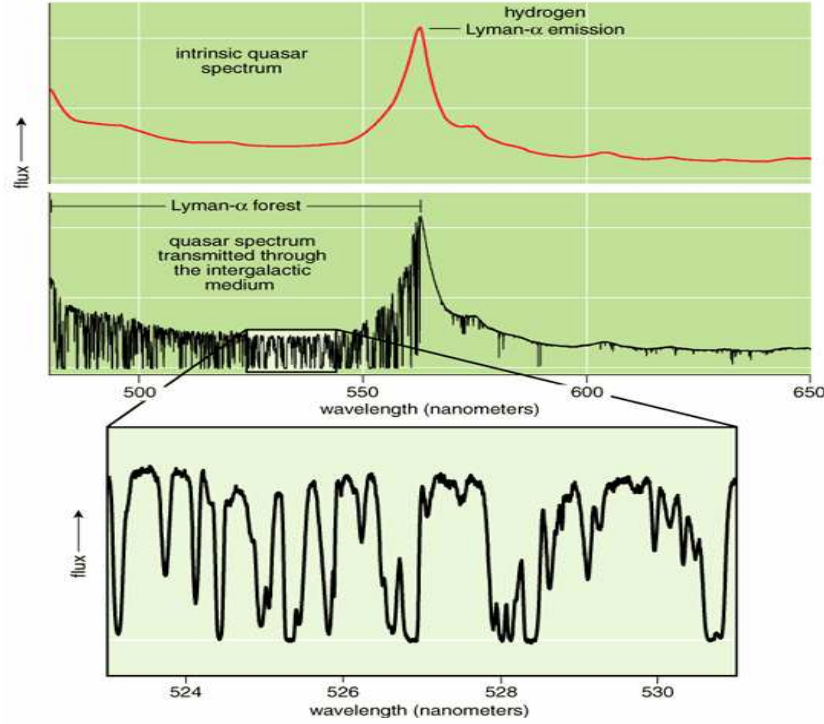


FIG. 5.1 – Spectre intrinsèque de quasar et spectre observé, avec un zoom sur la forêt Lyman- $\alpha$ . Issue de "The Cosmic Web", Robert A. Simcoe sur <http://www.americanscientist.org>

### 5.1.1 Forêt Lyman- $\alpha$ en observation

Le spectre émis par le quasar est représenté figure 5.1 et montre la différence entre la lumière émise et celle observée, après avoir traversé le milieu intergalactique, laissant apparaître la forêt Lyman- $\alpha$ . Il est possible de décomposer la forêt en différents systèmes d'absorption, associés à chaque raie et de déterminer via l'étude approfondie des raies les caractéristiques du système (température, vitesse...). Il s'agit de décomposer précisément les raies d'absorption selon un profil dit de Voigt (e.g. Carswell et al. 1987) et d'en déduire une description précise de la forme de la raie. Puis en ayant fait une analyse de toutes les raies constituant la forêt, étudier les propriétés statistiques de l'échantillon ainsi constitué. Il est possible d'explorer les caractéristiques du milieu intergalactique sans pour autant dépendre de l'identification ou du comptage des raies. Il suffit de mesurer la profondeur optique effective  $\tau_{eff}(z)$ , définie par le décrement en flux dans la forêt Lyman- $\alpha$  (expression simplifiée de l'équation 4.7) :

$$D_A = \left\langle 1 - \frac{f_{obs}}{f_{cont}} \right\rangle = 1 - e^{-\tau_{eff}} \quad (5.1)$$

avec  $f_{obs}$  le flux observé et  $f_{cont}$  le flux estimé du continu sans absorption. Avec des mesures de  $D_A$  disponibles sur un intervalle en décalage spectral, l'évolution de la forêt en fonction du décalage spectral peut être étudiée. Les incertitudes de cette mesure sont directement liées à notre méconnaissance du continu précis du

quasar, sur lequel est mesurée toute absorption. Quelques erreurs supplémentaires peuvent venir s'ajouter du fait de la contamination par certaines raies d'absorption métalliques, mais qui peuvent être dans la plupart des cas identifiées et enlevées de l'étude.  $\tau_{eff}$  a été mesurée par Press et al. (1993) pour un décalage spectral entre 2.5 et 4.5.

### 5.1.2 Regroupement des raies

Il est possible de mesurer la distribution spatiale de la matière en utilisant les nuages Lyman- $\alpha$  sur tout le domaine de décalage spectral que la forêt Lyman- $\alpha$  couvre. La mesure de la fonction de corrélation à deux points (que j'appellerai par la suite simplement fonction de corrélation) le long de plusieurs lignes de visée indépendantes, a été utilisée par Sargent et al. (1980) pour étudier le regroupement des absorbants responsables de la forêt Lyman- $\alpha$  et ils trouvent que celui-ci est moins prononcé que pour les galaxies. En effet, ils ne trouvèrent aucun signal entre 300 et 30 000  $km.s^{-1}$ . En outre, dû à la faible résolution de ces données, ils ne trouvèrent pas de signal aux petites échelles ( $\leq 300 km.s^{-1}$ ). Il fallut attendre les données à plus grande résolution pour détecter le signal (Muecket & Mueller 1987, Ostriker et al. 1988 et en particulier Cristiani et al. 1997), jusqu'à des séparations dans l'espace des vitesses de l'ordre de 100  $km.s^{-1}$ . En regroupant des données de suffisamment bonne précision sur assez de lignes de visée, il est possible de construire un échantillon ayant un chemin en décalage spectral suffisamment grand (e.g. McDonald et al. 2000). En plus du calcul de la fonction de corrélation pour détecter le regroupement des absorbants, il est possible d'étudier les structures que forment ces absorbants par d'autres moyens. Par exemple, Fang (1991) recherche les plus proches voisins dans un échantillon de raies d'absorption et détecte un signal sur des échelles de 30 à 50  $h^{-1}Mpc$ .

Les études sur les corrélations entre des raies d'absorption distinctes, en prenant en compte leurs caractéristiques (force relative de l'absorption, vitesse, etc...), donnent des informations importantes sur l'origine du signal de corrélation. Il est ainsi possible de rechercher des inhomogénéités dans la pression du gaz ou bien dans le flux ionisant le long de la ligne de visée (e.g. Zuo & Bond 1994). Ces inhomogénéités sont directement liées au regroupement des absorbants, par exemple il y a eu de longs débats sur la présence ou non de grands vides détectables dans le milieu intergalactique. Ces vides imposeraient un regroupement plus grand que celui observé. Néanmoins de grandes structures de l'ordre 30 à 40  $h^{-1}Mpc$ , présentant des sous-densités sont bien visibles, même si elles sont rares (Dobrzycki & Bechtold 1991, Cristiani et al. 1995, Rollinde et al. 2003).

## 5.2 Un milieu photo-ionisé

Les modèles de nucléosynthèse permettent de définir la densité d'hydrogène comme dépendant du paramètre de densité baryonique ( $\Omega_b$ ) et du taux d'expansion :

$$\bar{n}_H = (1.6 \times 10^{-7} \text{ cm}^{-3}) \left( \frac{\Omega_b h^2}{0.019} \right) (1+z)^3. \quad (5.2)$$

Comme l'hydrogène est prépondérant, la mesure de l'ionisation de cet élément est de toute première importance pour appréhender l'état physique du milieu intergalactique. Nous avons vu dans l'introduction que le milieu intergalactique a dû être réionisé par les premiers objets, sans savoir exactement de quoi il s'agit ; cela peut être une population d'étoiles très massives, des quasars etc... La réionisation a été réalisée par l'absorption de photons,  $H + \gamma \rightarrow p + e$ , avec une longueur d'onde plus courte que 912 Å ; c'est-à-dire avec une énergie supérieure à 13.6 eV, l'énergie de la limite de Lyman. La réionisation de l'hydrogène fut suivie par celle de l'hélium, et la fin de la réionisation correspond de façon assez schématique au moment où toutes les 'bulles' de réionisation se sont rejointes.

Le MIG est au final en équilibre de photo-ionisation, c'est-à-dire que d'une part on ne considère que le processus de ionisation par radiation, et on néglige l'ionisation par collision (car les densités sont trop faibles) ; d'autre part l'équilibre est réalisé lorsque le taux de photo-ionisation est parfaitement contrebalancé par le taux de recombinaison. On définit le taux de photo-ionisation par unité de volume pour chaque point d'un milieu optiquement mince composé d'hydrogène neutre comme étant :

$$n_{\text{HI}} \int_{\nu_L}^{\infty} \frac{4\pi J_{\nu} \sigma_{\text{H}}(\nu)}{h_P \nu} d\nu, \quad (5.3)$$

avec  $J_{\nu}$  l'intensité moyenne du flux ionisant (en unité d'énergie par unité de surface, de temps, d'angle solide et de fréquence) et  $h_P$  la constante de Planck. Et  $\sigma_H$  est la section efficace de photo-ionisation par des photons d'énergie supérieure à 13.6 eV, pour l'hydrogène au niveau fondamental qui peut s'approximer comme étant :

$$\sigma_{\text{H}}(\nu) = \sigma_L (\nu/\nu_L)^{-3}, \quad \sigma_L = 6.3 \times 10^{-18} \text{ cm}^2. \quad (5.4)$$

avec  $\nu$  fréquence du photon et  $\nu_L$  fréquence limite correspondant à la limite de Lyman. Et à l'équilibre cette ionisation est contrebalancée par le taux de recombinaison  $p + e \rightarrow H + \gamma$  par unité de volume :

$$n_e n_p \alpha_A(T) \quad (5.5)$$

avec  $n_e$  et  $n_p$  les densités d'électrons et de protons respectivement et  $\alpha_A$  est le coefficient de recombinaison, c'est-à-dire le produit de la section efficace de capture d'un électron avec la vitesse d'un électron, moyennée sur une distribution thermique et sommée sur tous les niveaux atomiques.

Le milieu intergalactique est complètement réionisé lorsque un photon ionisant a été émis pour chaque atome d'hydrogène et lorsque le taux d'émission des photons contrebalance le taux de recombinaison. Si cette ionisation était complète elle se traduirait par l'absence d'absorptions dans les spectres de galaxies et de quasars à grand décalage spectral.

### 5.3 Modèle actuel

Les simulations sont le meilleur moyen à notre portée pour confronter les modèles aux observations des spectres. En effet, elles offrent la possibilité de



traiter à la fois l'évolution de la matière noire et du gaz en partant des perturbations primordiales et en plus il est possible d'en extraire des spectres simulés afin de construire des échantillons comparables aux observations.

### 5.3.1 La forêt Lyman- $\alpha$

L'interprétation des absorptions observées dans la forêt Lyman- $\alpha$  a été donnée par Sargent et al. (1980), qui démontrèrent que celles-ci provenaient d'un milieu intergalactique. Le modèle qui permet d'expliquer ces absorptions remonte à Gunn & Peterson (1965) qui ont développé l'idée d'un milieu intergalactique hautement ionisé, ponctué par des nuages d'hydrogène neutre, plus tard leurs noms furent associés à l'effet Gunn-Peterson <sup>1</sup>. Ce modèle de nuages confinés (voir Rauch 1998 pour une revue de ces modèles) soit par la pression d'un milieu intergalactique chaud (Sargent et al. 1980) soit par gravité (Melott 1980) soit par des mini-halos de matière noire froide (Rees 1986), ou par une combinaison du tout (Petitjean et al. 1993), a été très utilisé jusqu'à ce que l'on se rende compte que la confrontation des modèles aux observations posait d'évidents problèmes d'adaptation. Il fallut attendre les observations à très grande résolution et des simulations hydrodynamiques pour faire évoluer ce modèle vers une conception d'ensemble plus cohérente.

Les simulations ont montré qu'il pouvait exister des filaments de matière reliant les structures les plus denses dans l'univers, la 'toile cosmique'. Elles ont aussi montré que cette toile devait contenir du gaz ionisé par le flux issu des sources tels que les quasars. Ces simulations ont permis de reproduire de façon remarquable les absorptions en hydrogène neutre observées le long des lignes de visée (e.g. Cen et al. 1994, Zhang et al. 1995). En particulier la distribution en densité de colonne (c'est-à-dire le nombre de systèmes par unité de colonne densité et par unité de décalage spectral) de ces modèles reproduit la loi de puissance des observations (Hernquist et al. 1996).

De ces résultats est apparu un nouveau paradigme, les absorptions responsables de la forêt Lyman- $\alpha$  sont issues d'hydrogène neutre présent dans des régions denses ou peu denses de l'univers. Il s'agit donc d'une structure filamentaire représentée par des régions denses, séparés par de grands 'vides', le tout étant fortement ionisé. La figure 5.2 montre une simulation réalisée par Cen, où l'on peut comparer ce qu'était le milieu intergalactique il y a douze milliards d'années et ce qu'il est aujourd'hui.

Le gaz responsable de ces absorptions doit avoir sa température directement reliée à sa densité. Son état est en équilibre entre le refroidissement adiabatique dû à l'expansion et la photoionisation par le fond ultra-violet (Hui & Gnedin 1997). Les équations régissant l'état du milieu peuvent être contraintes de façon observationnelle (Schaye et al. 2000, Choudhury et al. 2001). Ainsi une densité donnée correspond à une température donnée, qui elle-même correspond à une fraction de gaz ionisé et qui finalement donne la profondeur optique en hydro-

---

<sup>1</sup>Dans un univers en expansion, rempli majoritairement d'hydrogène neutre, l'opacité du milieu s'étend à toute la partie bleue du spectre à partir de la transition Lyman- $\alpha$ . Une absorption continue doit donc être détectée au niveau du continu du quasar dans l'ultra-violet (laboratoire) ; c'est ce qui est appelé l'effet Gunn-Peterson

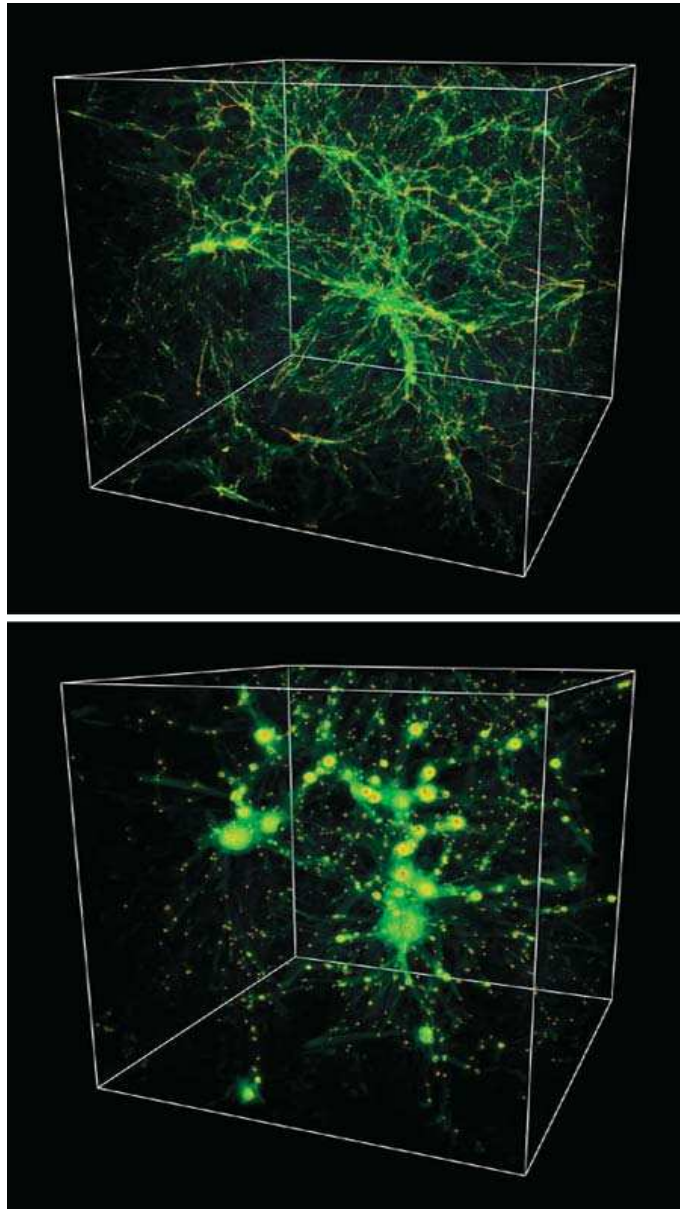


FIG. 5.2 – En haut, le milieu intergalactique tel qu'il était il y a 12 milliards d'années, en bas tel qu'il est aujourd'hui. La boîte du bas est quatre fois plus large que celle du haut qui fait 30 millions d'années lumière de côté. Issue de "The Cosmic Web", Robert A. Simcoe sur <http://www.americanscientist.org>

gène neutre. Les simulations ont ainsi mené à formuler l'approximation de l'effet Gunn-Peterson fluctuant, (Croft et al. 1998) où la profondeur optique en hydrogène neutre,  $\tau_{HI}$ , est reliée à la densité de la matière noire ou du gaz (puisque'il se suivent mutuellement dans ce milieu de faible densité où la gravitation domine),  $\rho : \tau_{HI} \propto \rho^{1.6}$ .

### 5.3.2 L'enrichissement en métaux du milieu intergalactique

La spectroscopie haute résolution sur les télescopes de classe 8-10 m a permis d'étudier les absorptions métalliques dans les régions de faible densité de l'univers. Par exemple Songaila & Cowie (1996) trouvent que la plupart des systèmes Lyman- $\alpha$  à un décalage spectral de 3 avec une densité de colonne  $N_{HI} > 10^{15} \text{ cm}^{-2}$  et que la moitié des systèmes avec  $N_{HI} > 10^{14.5} \text{ cm}^{-2}$  ont une absorption en C IV qui leur est associée. Ainsi par certains mécanismes les métaux sont transportés sur des distances cosmologiques. Pour identifier de tels mécanismes il faut examiner la dépendance de la métallicité en fonction de la densité : un gradient fort tendrait à prouver que l'enrichissement est local tandis qu'une distribution uniforme suggérerait un enrichissement à une époque très lointaine (cf. Rollinde et al. 2001 et Aracil et al. 2004).

Aguirre et al. (2002) ont étudié la métallicité en C IV dans des régions de très faible densité en utilisant des spectres à très haute résolution de quasars à un décalage spectral entre 2 et 4. Ils trouvent d'un côté un faible gradient de densité ainsi qu'une évolution faible de la métallicité du milieu intergalactique sur l'intervalle en décalage spectral étudié. Ce manque d'évolution a aussi été trouvé par Songaila (2001) pour des décalages spectraux entre 2 et 5.5. Aguirre et al. (2003) ont étendu leur recherche au Si IV, qui trace des densités plus grandes que le C IV, et ont confirmé cette 'non-évolution'. Ces résultats pointent donc dans la direction d'un enrichissement bien avant un décalage spectral de 3. D'un autre côté, Pettini et al. (2002) ont observé des vents émergeant des galaxies à brisure de Lyman <sup>2</sup> ( $z \sim 3$ ), ce qui pourrait être un des mécanismes d'enrichissement local. Adelberger et al. (2003) trouvent aussi un enrichissement fort dans les lignes de visée passant à proximité de galaxies, ce qui est une preuve d'un enrichissement *in situ* du milieu. Ainsi les mécanismes et l'époque de l'enrichissement du milieu intergalactique restent encore incertains. La réponse viendra une nouvelle fois de la confrontation des simulations aux observations. Par exemple Oppenheimer & Davé (2006) testent les modèles d'enrichissement grâce à des simulations hydrodynamiques de formation des structures, en y incorporant des jets provenant des régions de formation d'étoiles dans les galaxies. Il ressort de cette étude que les vents issus des galaxies à forte formation d'étoiles peuvent produire suffisamment d'absorbants en C IV à grand décalage spectral et leur nombre est relativement invariant d'un décalage spectral de 5.5 à 1.5, ce qui correspond en partie aux observations.

## 5.4 Tomographie du milieu

Ce qui manque à l'étude du regroupement des raies d'absorption ce sont des informations en trois dimensions. En effet, pour comprendre la géométrie des absorbants et dissocier leur position en vitesse de leur position dans l'espace il est nécessaire d'utiliser des informations transverses, c'est à dire perpendiculaires à la ligne de visée. Les premiers pas dans ce sens pour estimer la taille

---

<sup>2</sup>Appelées aussi galaxies 'starburst' à grand décalage spectral, approximativement entre  $2.5 < z < 5$ . Elles sont identifiées par les couleurs de leur DSE autour de la discontinuité de Lyman à 912 Å

des absorbants ont été réalisés en utilisant des paires de quasars projetées dans le ciel. Il est envisageable de pousser cette recherche d'information à la reconstruction en trois dimensions du milieu intergalactique, en ayant recours à des groupes de quasars.

#### 5.4.1 Utilisation de paires de quasars

La première utilisation de paires de quasars s'est faite avec des images de quasars subissant l'effet de lentille gravitationnelle. Les images ne sont généralement pas séparées de plus de quelques arcsecondes, et donnent donc des informations sur les absorbants sur une échelle inférieure à 100 kpc. L'une des observations les plus remarquables a été réalisée par Smette et al. (1995) sur le quasar HE 1104-1805A,B. Elles donnèrent une limite inférieure sur la taille des absorbants de la forêt Lyman- $\alpha$  de  $25 h^{-1} \text{ kpc}$ . Les quasars soumis à l'effet de lentille peuvent aussi être utilisés pour étudier les absorptions métalliques à grand décalage spectral et leur lien avec les galaxies (cf. Smette et al. 1995).

On utilise aussi les paires formées de deux quasars proches, ceux-ci ont généralement des séparations de l'ordre de la minute d'arc ou plus dans le ciel. Cette plus grande séparation permet de sonder une plus grande échelle de cohérence. La paire Q1343+266A,B séparée par 9.5 arcseconde permet de montrer que les absorbants s'étendent sur une échelle de plus d'une centaine de kilo parsecs (Bechtold et al. 1994, Dinshaw et al. 1994); Fang et al. 1996 proposèrent pour la même paire une taille de  $500 h^{-1} \text{ kpc}$ . Différentes études tendent à montrer que les absorbants peuvent s'étendre sur des échelles de l'ordre du méga parsec et plus (e.g. Petitjean et al. 1998, D'Odorico et al. 1998, Young et al. 2001, Aracil et al. 2002, Rollinde et al. 2003). L'utilisation de paires permet aussi de mettre en évidence des structures à différentes échelles comme les zones de sous-densité (Rollinde et al. 2003), ou d'étudier l'effet de l'ionisation des quasars sur leur environnement. S'il est possible de trouver des paires de quasars, on s'attend aussi à pouvoir utiliser des regroupements de lignes de visée afin de sonder le milieu intergalactique en trois dimensions.

#### 5.4.2 Futur de la tomographie - les groupes de quasars

La reconstruction des propriétés du milieu intergalactique à travers l'utilisation de lignes de visée proches est présentée en détail dans Nusser & Haehnelt (1998) et Pichon et al. (2001). Je vais présenter ici une partie des résultats de Pichon et al. (2001), qui montrent toute la complexité de la mise en oeuvre de telles méthodes.

Pichon et al. (2001) proposent l'implémentation d'une méthode d'inversion bayésienne afin de retrouver les propriétés du milieu intergalactique à partir d'observations de la forêt Lyman- $\alpha$  dans des spectres de quasars à grand décalage spectral. La physique du gaz est relativement simple et pour résumer ce qui a été dit précédemment : "son état thermique est gouverné par le réchauffement via la photo-ionisation et par le refroidissement adiabatique. Sa dynamique résulte d'effets de gravité sur de grandes échelles et de pression sur de petites échelles. La matière noire et les baryons se suivent mutuellement et la forêt Lyman- $\alpha$

est due à des variations de densité dans le milieu intergalactique. Le gaz semble être distribué le long de filament et/ou de feuilles d'une certaine extension." (Pichon et al. 2001). La précision de ces études qui ont mené à ces conclusions sur la physique du gaz peut être améliorée en augmentant le nombre de lignes de visée observées à grand décalage spectral et à haute résolution. Une autre approche peut être considérée, qui consiste à utiliser des groupes de quasars afin de contraindre la distribution en trois dimensions du gaz dans le même champ, l'idéal pour cela étant d'accroître au maximum la densité de lignes de visée observées. La reconstruction a pour but de retrouver les structures (filaments, feuilles, ...) et les propriétés physiques du gaz (température, vitesse particulière). En se basant sur des a priori sur la fonction de corrélation, on montre comment recouvrir la topologie en trois dimensions des grandes structures en utilisant un réseau de lignes de visée adjacentes.

L'idée de base est d'interpoler entre des lignes de visée proches les champs qui sont mesurés le long des lignes de visée. Ceci nécessite dans un premier temps de faire des suppositions sur la nature des champs, ce qui est assez risqué car, "pour autant que nous en sachions l'espace entre les lignes de visée pourrait bien être vide" (Pichon et al. 2001). Il faut donc trouver une solution unique à l'inversion de l'équation 5.6, ce qui nécessite plusieurs hypothèses, notamment sur le paramètre Doppler et sur le champ des vitesses particulières.

La profondeur optique,  $\tau_\ell(w)$ , le long d'une ligne de visée,  $\ell$ , à une position projetée  $\mathbf{x}_{\perp,\ell} \equiv (y_\ell, z_\ell)$  sur le ciel, et dans l'espace des vitesses,  $w$ , est reliée à la densité d'hydrogène neutre,  $n_{\text{HI}}$ , par (cf. équation 23.11 de Peebles : "Principles of physical cosmology" et équation 1 de Pichon et al. 2001) :

$$\begin{aligned} \tau_\ell(w) = & \frac{c \sigma_0}{H(\bar{z})\sqrt{\pi}} \iint \left( \int_{-\infty}^{+\infty} \frac{n_{\text{HI}}(x, \mathbf{x}_\perp)}{b(x, \mathbf{x}_\perp)} \right. \\ & \left. \exp \left( - \frac{(w - x - v_p(x, \mathbf{x}_\perp))^2}{b(x, \mathbf{x}_\perp)^2} \right) dx \right) \\ & \delta_D(\mathbf{x}_\perp - \mathbf{x}_{\perp,\ell}) d^2\mathbf{x}_\perp, \quad \ell = 1 \dots L, \end{aligned} \quad (5.6)$$

avec  $\sigma_0$  la section efficace effective intégrée à travers la résonance dans la transition Lyman- $\alpha$  1s à 2p de l'atome d'hydrogène,  $H(\bar{z})$  la constante de Hubble à un décalage spectral moyen  $\bar{z}$  et  $v_p(x)$  est la projection de la vitesse particulière le long de la ligne de visée. La double intégrale sur  $\mathbf{x}_\perp$  correspond à l'intégration dans les directions perpendiculaires aux lignes de visée.  $\delta_D$  est la distribution 2D de Dirac. Le paramètre Doppler  $b(\mathbf{x})$  est considéré comme fonction de la température locale du milieu intergalactique au point  $\mathbf{x} \equiv (x, \mathbf{x}_\perp)$  où  $x$  est la coordonnée dans l'espace réel exprimée en km/s [=  $rH(\bar{z})$ ]. Il s'agit d'inverser  $n_{\text{HI}}(x, \mathbf{x}_\perp)$ ,  $b(x, \mathbf{x}_\perp)$  et  $v_p(x, \mathbf{x}_\perp)$ , à partir de la connaissance d'un échantillon de ligne de visée,  $\ell = 1 \dots L$ .

Pichon et al. (2001) testent en particulier cette méthode sur des simulations à N-corps. Ils étudient notamment les effets selon le nombre et la séparation des lignes de visée et estiment que plusieurs lignes de visée à faible résolution peuvent être utilisées pour améliorer la reconstruction en vitesse d'une ligne de visée à haute résolution.

La figure 5.3 montre les résultats de différentes reconstructions. On peut voir que pour une faible séparation (2.5 Mpc) la reconstruction est très bonne, et que pour des séparations supérieures on perd les informations à petite échelle, mais les grandes structures sont bien recouvrées. La topographie des structures est bien définie. Par contre, comme attendu, la densité n'est pas bien retrouvée pour les séparations les plus larges.

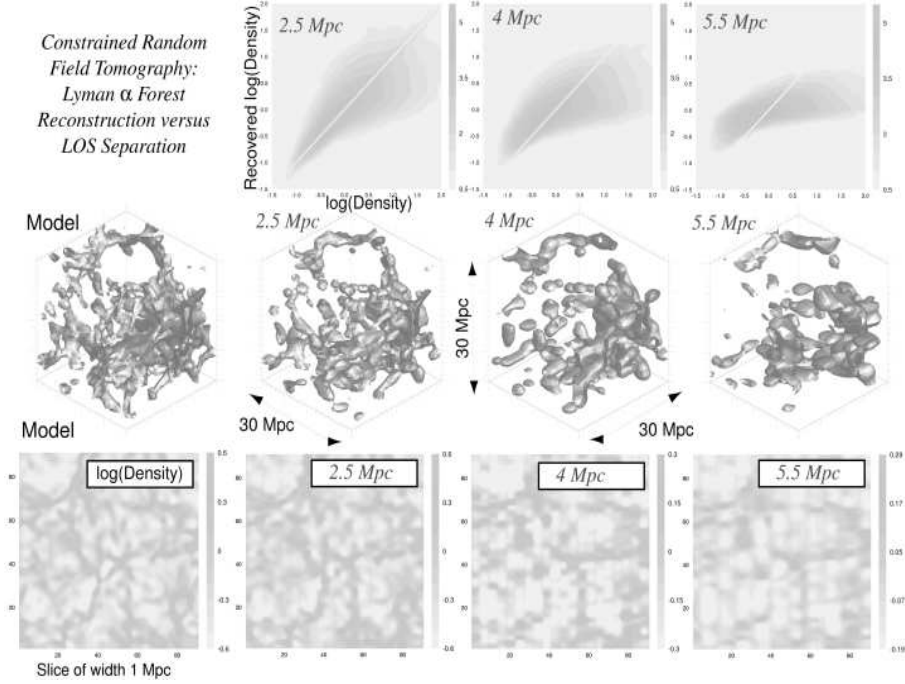


FIG. 5.3 – Panneau de haut en bas et de gauche à droite : la densité recalculée en fonction du champ de densité initial, en fonction de la distance entre les lignes de visée (2.5 Mpc, 4 Mpc et 5.5 Mpc). Le modèle et les différentes reconstructions selon la séparation des lignes de visée. Coupe de la simulation et des champs reconstruits, l'échelle est en pixels. Issue de Pichon et al. 2001

Il faut se soucier du nombre de lignes de visée à employer. La figure 5.4 représente le facteur de qualité qui mesure l'inverse du résidu de "mauvaise adaptation" en unité de variance pour la vitesse. Ce facteur de qualité,  $Q$ , augmente avec le nombre de lignes de visée mises en jeu, ce qui est attendu. Par contre  $Q$  augmente aussi avec la distance entre les lignes de visée jusqu'à atteindre une saturation. Ceci vient du fait que le groupe de lignes de visée permet de contraindre les vitesses transverses que sur des échelles intermédiaires. C'est-à-dire qu'il faudrait des lignes de visée extrêmement proches (distance inférieure à 1 Mpc) pour sonder les petites échelles. D'un autre côté les grandes structures ne peuvent être correctement sondées qu'avec des lignes de visée suffisamment séparées, jusqu'à 20 Mpc. La compétition entre ces deux effets (petites et grandes échelles) permet de fixer une séparation optimale des lignes de visée en fonction de leur nombre, d'après la figure 5.4 la séparation optimale pour un échantillon de lignes de visée de  $11 \times 11$  est de 5 Mpc.

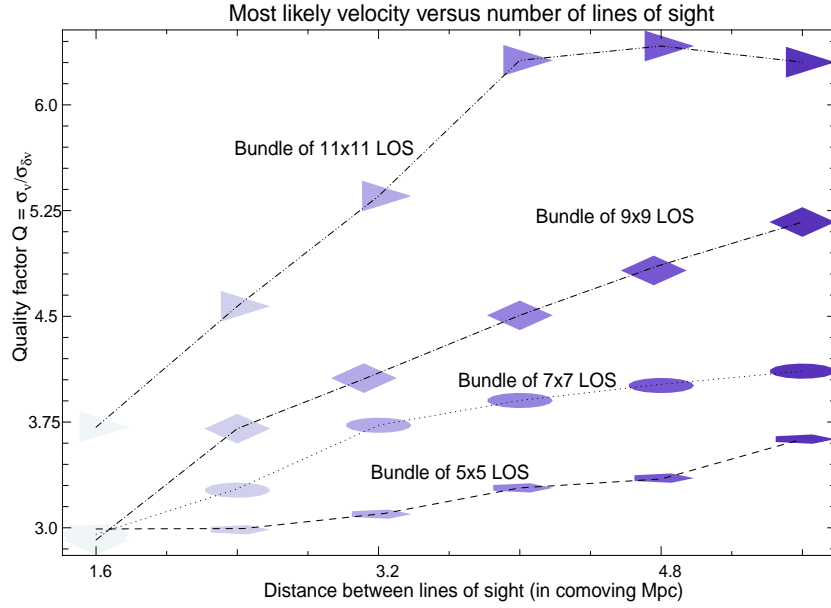


FIG. 5.4 – Qualité de la reconstruction en fonction de la séparation et du nombre de lignes de visée. Augmenter l'échantillonnage sur le ciel diminue la dispersion entre la vitesse recouvrée et la vitesse mesurée. Noter la saturation à 5 Mpc pour un groupe de lignes de visée de 11. Issue de Pichon et al. 2001

Un point important dans cette analyse est que comme la séparation pour un nombre donné de lignes de visée est plutôt grande, les informations à petites échelles dans la reconstruction ne sont issues que d'une seule ligne de visée. Il suffit donc d'avoir un spectre à haute résolution avec un échantillon de spectres à basse résolution séparés par des distances plus petites ou de l'ordre de 4-5 Mpc. Le spectre haute résolution doit se situer au centre du groupe des lignes de visée.

La recherche et l'utilisation des groupes de quasars afin de réaliser une inversion du milieu intergalactique dans le but de cartographier l'Univers n'est pas une entreprise aisée. Il faut prendre en compte tous les biais lors du calcul de l'inversion, les effets à petites échelles que l'on n'appréhende pas encore complètement, ainsi que tous les effets systématiques lors de l'analyse des observations. Ceci implique qu'il y a beaucoup à apprendre des échantillons de paires de quasars ainsi que des simulations dont nous disposons déjà. C'est dans cette optique qu'a été réalisé le travail présenté en section 5.5.

## 5.5 Transverse and longitudinal correlation functions in the Intergalactic Medium from 32 close pairs of high-redshift quasars<sup>3</sup>

F. Coppolani<sup>1</sup>, P. Petitjean<sup>2,3</sup>, F. Stoehr<sup>2</sup>, E. Rollinde<sup>4,2</sup>, C. Pichon<sup>2</sup>, S. Colombi<sup>2</sup>, M.G. Haehnelt<sup>5</sup>, B. Carswell<sup>5</sup>, R. Teyssier<sup>6</sup>

<sup>1</sup> European Southern Observatory, Alonso de Córdova 3107, Casilla 19001, Vitacura, Santiago, Chile - email : fcoppola@eso.org

<sup>2</sup> Institut d'Astrophysique de Paris, UMR 7095 CNRS & Université Pierre et Marie Curie, 98 bis boulevard d'Arago, 75014 Paris, France

<sup>3</sup> LERMA, Observatoire de Paris, 61, avenue de l'observatoire F-75014 Paris, France

<sup>4</sup> IUCAA, Post Bag 4, Ganesh Khind, Pune 411 007, India

<sup>5</sup> Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK

<sup>6</sup> DAPNIA-CEA, Saclay

**abstract** We present the transverse flux correlation function of the Lyman- $\alpha$  forest in quasar absorption spectra at  $z \sim 2.1$  from VLT-FORS and VLT-UVES observations of a total of 32 pairs of quasars; 26 pairs with separations in the range  $0.6 < \theta < 4$  arcmin and 6 pairs with  $4 < \theta < 10$  arcmin. Correlation is detected at the  $3\sigma$  level up to separations of the order of  $\sim 4$  arcmin (or  $\sim 4.4h^{-1}$  Mpc comoving at  $z = 2.1$  for  $\Omega_m = 0.3$  and  $\Omega_\Lambda = 0.7$ ). We have, furthermore, measured the longitudinal correlation function at a somewhat higher mean redshift ( $z = 2.39$ ) from 20 lines of sight observed with high spectral resolution and high signal-to-noise ratio with VLT-UVES. We compare the observed transverse and longitudinal correlation functions to that obtained from numerical simulations and illustrate the effect of spectral resolution, thermal broadening and peculiar motions. The shape and correlation length of the correlation functions are in good agreement with those expected from absorption by the filamentary and sheet-like structures in the photoionized warm intergalactic medium predicted in CDM-like models for structures formation. Using a sample of 139 C IV systems detected along the lines of sight toward the pairs of quasars we also investigate the transverse correlation of metals on the same scales. The observed transverse correlation function of intervening C IV absorption systems is consistent with that of a randomly distributed population of absorbers. This is likely due to the small number of pairs with separation less

---

<sup>3</sup>Based on observations carried out at the European Southern Observatory with UVES (ESO program No. 65.O-299 and the Large Program 'The Cosmic Evolution of the IGM' 166.A-0106), FORS2 (ESO program No. 66.A-0183) and FORS1 (ESO programs No. 69.A-0457 and 70.A-0032) on the 8.2 m VLT telescopes Antu, Kuyen and Melipal operated at Paranal Observatory; Chile.



than 2 arcmin. We detect, however, a significant overdensity of systems in the sightlines towards the quartet Q 0103–294A&B, Q 0102–2931 and Q 0102–293 which extends over the redshift range  $1.5 \leq z \leq 2.2$  and an angular scale larger than 10 arcmin.

**keywords :**

*Methods* : data analysis - N-body simulations - statistical -

*Galaxies* : intergalactic medium - quasars : absorption lines -

*Cosmology* : dark matter

### 5.5.1 Introduction

The numerous H I absorption lines seen in the spectra of distant quasars, the so-called Lyman- $\alpha$  forest, contains precious information on the spatial distribution of neutral hydrogen in the Universe. Unravelling this information from individual spectra has for a long time proven difficult and ambiguous (see Rauch 1998 for a review). Studies of the correlation of the Lyman- $\alpha$  forests observed in the two spectra of QSO pairs have been instrumental in measuring the spatial extent of absorbing structures. The Lyman- $\alpha$  forests in the spectra of multiple images of lensed quasars or pairs of quasars with separations of a few arcsec (Bechtold et al. 1994 ; Dinshaw et al. 1994 ; Smette et al. 1995 ; Impey et al. 1996 ; Rauch, Sargent & Barlow 1999 ; Becker, Sargent & Rauch 2004) appear nearly identical implying that the absorbing structures have sizes  $>50h_{70}^{-1}$  kpc. Significant correlation between absorption spectra of adjacent lines of sight toward quasars still exists for separations of a few to ten arcmin suggesting a size or better a coherence length of the structures larger than  $500 h_{70}^{-1}$  kpc (e.g. Shaver & Robertson 1983 ; Dinshaw et al. 1997 ; Petitjean et al. 1998 ; D’Odorico et al. 1998 ; Crotts & Fang 1999 ; Young, Impey & Foltz 2001 ; Aracil et al. 2002, Rollinde et al. 2003) and a non-spherical geometry of the absorbing structures (Rauch & Haehnelt 1995 ; Rauch et al. 2005). On even larger scales, Williger et al. (2000) still find evidence for an excess of clustering on 10 Mpc scales.

Numerical simulations of the warm photoionized Intergalactic Medium within the framework of cold dark matter (CDM) like models of structure formation have demonstrated that the neutral gas density traces the underlying dark matter density field on scales larger than the Jeans length of the gas (e.g. Cen et al. 1994 ; Petitjean, Mückel & Kates 1995 ; Theuns et al. 1998). The picture of the Lyman- $\alpha$  forest arising from the density fluctuations in a warm photoionized Intergalactic Medium distributed as expected in a CDM model explains the statistical properties of individual QSO absorption spectra very well (see Weinberg et al. 1999 for a review). Most of the baryons are located in filaments and sheets which are only overdense by factors of a few and produce absorption in the column density range  $10^{14} < N_{\text{HI}} < 10^{15} \text{ cm}^{-2}$  at  $z \sim 2$ . On the other hand, most of the volume is occupied by underdense regions that produce absorption with  $N_{\text{HI}} < 10^{14} \text{ cm}^{-2}$ . Analytical calculations and numerical simulations of the spatial distribution of neutral hydrogen in  $\Lambda$ CDM models are also able to reproduce the large observed transverse correlation length of the Lyman- $\alpha$  forest in the absorption spectra of QSO pairs (Bi 1993 ; Miralda-Escudé et al. 1996 ;

Charlton et al. 1997; Viel et al. 2002; Rollinde et al. 2003; Rauch et al. 2005).

As pointed out by several authors a comparison of the transverse correlation to the correlation observed along the line of sight, can be used to carry out a variant of the Alcock & Paczyński (1979) test to put constraints on the geometry of the Universe (Hui, Stebbins & Burleset 1999; Mc Donald & Miralda-Escudé 1999; Mc Donald 2003). This provides strong motivation for an (accurate) measurement of the transverse correlation function.

In a previous work we have used 5 pairs and a group of 10 quasars with separations in the range 1-10 arcmin to investigate whether the longitudinal and transverse correlation functions were consistent with those expected in  $\Lambda$ CDM models (Rollinde et al. 2003). We reported a somewhat marginal detection of a transverse correlation up to separations of 3 to 4 arcmin. We have assembled here a significantly larger sample of 32 QSO pairs. The new sample consists of 26 pairs with separations in the range 0.6–4 arcmin (corresponding to  $\simeq 0.2$  to  $1.4 h^{-1}$  Mpc proper at  $z=2.1$  for  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$ ) and 6 pairs with separations in the range 5–10 arcmin.

Details of the observations and simulations are given in Sections 5.5.2 and 5.5.3, respectively. In Section 4 we define and discuss our measurements of the observed longitudinal and transverse flux correlation functions. Section 5 compares the observed and simulated correlation functions. We investigate the transverse correlation of C IV absorption systems in Section 6. Our conclusions are given in Section 7. Comments on individual lines of sight are given in Appendix A. Metal-line lists and QSO spectra are given in Appendix B and C available in the electronic version of the paper.

### 5.5.2 Observations

The first release of the 2dF quasar survey has significantly enlarged the number of known quasar pairs with arcmin separation (Outram, Hoyle & Shanks 2001). We have selected pairs with the following criteria to enlarge the number of small separation pairs with respect to the sample of Rollinde et al. (2003): (i) the separation of the two quasars should be in the range 1–4 arcmin where the correlation is expected and observed to be strong; (ii) the quasars should be brighter than  $m_V = 20.30$  to keep observing time in reasonable limits; (iii) the emission redshifts of the two quasars should be larger than  $z \sim 2.1$  to increase the wavelength range over which high S/N ratio can be obtained (FORS is not sufficiently sensitive below 3500 Å); (iv) the redshift difference should be smaller than  $\Delta z \sim 0.5$  (for most of them 0.3) to maximise the wavelength range over which correlation can be studied.

There are 22 quasar pairs in the 2dF survey which meet our criteria of which we observed 20. We have observed two additional pairs not contained in 2dF: J 123510.5-010746–J 123511.0-010830 with a separation of 0.74 arcmin and Q 1207-1057–Q 1206-1056 with a separation of 3.5 arcmin. The spectra were obtained with FORS1 and 2 mounted on the VLT-UT2 and UT3 telescopes of ESO using the grism GR630B and a 0.7 arcsec slit. The spectra were reduced using standard procedures available in the context LONG of the ESO data reduction package MIDAS. Master bias and flat-fields were produced using

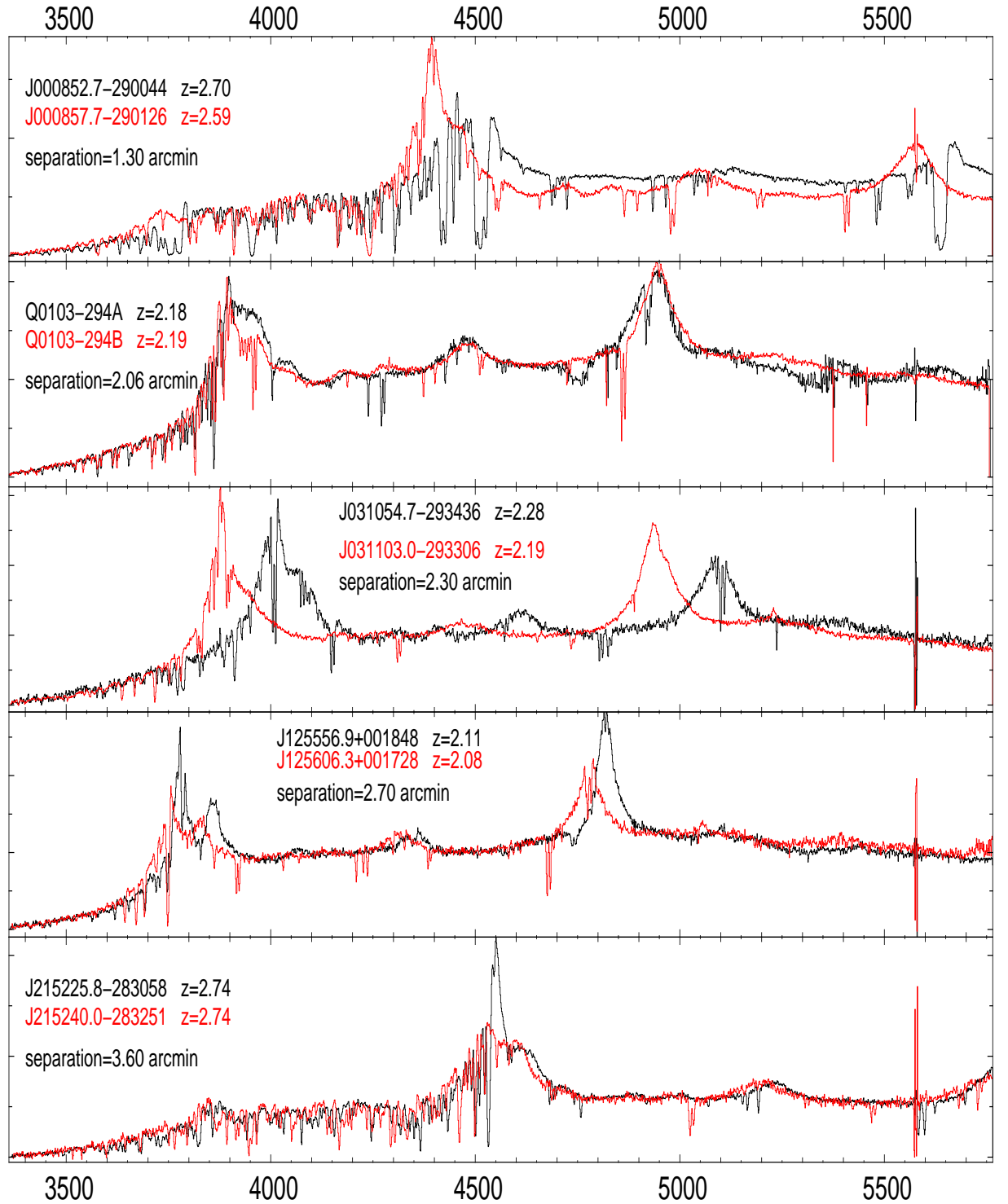


FIG. 5.5 – Typical observed spectra of QSO pairs in order of increasing right ascension. QSO names, emission redshifts and angular separations are given in the top-left corner of each sub-panel. Other spectra are presented in Appendix C published in the electronic version of the paper.

day-time calibrations. Bias subtraction and flat-field division were performed on science and calibration images. A correction for 2D distortion was applied. The sky level was evaluated in two windows on both sides of the object offset along the slit direction and subtracted on the fly during the optimal extraction of the object. The spectra were then wavelength calibrated over the range  $3400 < \lambda < 6000 \text{ \AA}$ . The final pixel size is  $1.18 \text{ \AA}$  corresponding to a resolution of  $R = 1400$  or  $\text{FWHM} = 220 \text{ km s}^{-1}$  at  $3800 \text{ \AA}$ . The exposure times have been adjusted in order to obtain a typical signal-to-noise ratio of  $\sim 10$  at  $3500 \text{ \AA}$ . The sharp decrease of the detector sensitivity below  $4000 \text{ \AA}$  prevents scientific analysis below  $\sim 3500 \text{ \AA}$ . At  $\lambda \sim 4500 \text{ \AA}$  the S/N ratio is usually larger than 70. The final sample consists of 58 QSOs (44 QSOs new from this program, 12 QSOs from Rollinde et al. 2003 plus the pair UM680-UM681 presented in D'Odorico et al. 2002). The total number of pairs included in our analysis of the transverse correlation function is 32 somewhat larger than half the total number of QSOs due to the additional pairings of Q 0103–294A&B, Q 0102–2931 and Q 0102–293 which form a group (alternative names are, respectively, J010534.7-290917, J010538.3-291106, J010518.0-291510 and J010502.8-290618). One further pair in the sample (J 123510.5-010746–J 123511.0-010830) was at the end not included in our analysis of the transverse correlation function as the redshift overlap of the Lyman- $\alpha$  forest is too small to contribute in a statistically significant way.

Table 5.1 gives a summary of the sample including emission redshifts, angular separation of the pairs on the sky, the mean S/N ratio in the wavelength range of interest. Emission redshifts were determined by fitting a Gaussian function to the C IV emission line when present in the spectrum or to the Lyman- $\alpha$  emission line otherwise. Typical spectra are shown in Fig 5.5. Other spectra are presented in Appendix C published in the electronic version of the paper.

The spectra have been normalized using a spline fit to the continuum. This operation is important as it can affect the estimate of the mean flux which is in turn critical for the flux correlation function estimate. In our sample most of the Lyman- $\alpha$  forest common to two QSOs lies at  $z < 2.3$ . At this redshift the density of absorption features is moderate and there are numerous spectral regions with no absorption. This is why continuum fitting is reliable. For the same reason, the identification of metal absorption systems in the Lyman- $\alpha$  forest region of the spectrum is unproblematic. We have checked that the mean absorption of the spectra in our sample is consistent with that measured from the data of the VLT Large Program (LP) 'The Cosmic Evolution of the IGM' (Aracil et al. 2004, 2006 and below).

To calculate the longitudinal correlation function we also use the data from the LP 'The Cosmic Evolution of the IGM' which has produced a sample of absorption spectra of homogeneous quality suitable for studying the Lyman- $\alpha$  forest in the redshift range 1.7–4.5. The spectra of the LP have been taken with VLT-UVES and have high resolution ( $R \sim 45000$ ), high signal to noise ratio (30 and 60 per pixel at respectively 3500 and 6000  $\text{\AA}$ ) and cover the wavelength ranges 3100–5400 and 5450–9000  $\text{\AA}$ . Details of the data reduction and normalization of the spectra are given in Aracil et al. (2004, 2006).

### 5.5.3 Numerical simulations

In this paper we use two numerical simulations to estimate the errors and the effect of numerical resolution, redshift distortion and the thermal state of the gas on the correlation functions : a large size dark-matter only simulation and a smaller size full hydrodynamical simulation. For both simulations we assume parameters consistent with the fiducial concordance cosmological model  $\Omega_\Lambda = 0.7$ ,  $\Omega_m = 0.3$ . Hubble constants are  $h = 0.65$  and  $h = 0.7$  and normalization of the fluctuation amplitude of the matter power spectrum are  $\sigma_8 = 0.93$  and  $\sigma_8 = 1.0$  for the dark-matter only and the full hydrodynamical simulation, respectively. The assumed baryon density in the hydro-dynamical simulations is  $\Omega_b = 0.04$ . The simulations were performed on the computers of the Institut du Developpement et des Ressources en Informatique Scientifique (IDRIS) in Orsay.

The hydro-dynamical simulation of 40 Mpc box-size and dark-matter only simulation of  $100 h^{-1}\text{Mpc}$  box-size are used to catch both the statistical aspects and the effect of gas physics. In the hydro-simulation some weak bias on the correlation functions due to the box-size is expected at separation larger than 4 Mpc. However, larger hydrodynamical simulations that still resolve the Jeans length at least marginally are currently not feasible.

The dark-matter only simulation was performed with the Particle - Mesh (PM) code described in Pichon et al. (2001) and was also used in Rollinde et al. (2003). The simulation has 16 million particles and a box-size of  $100 h^{-1}\text{Mpc}$ .<sup>4</sup> The large box-size ensures a sufficient statistical sampling on large scales where thermal effects and pressure effects which are modelled approximately in the dark-matter only simulation are less important. Initial conditions were set up using a standard CDM transfer function (Efstathiou, Bond & White 1992). To construct mock Lyman- $\alpha$  spectra from the simulated data, we proceed as in Rollinde et al. (2001, 2003) applying simple semi-analytical prescriptions taking into account thermal broadening and redshift distortion. The density and velocity fields are interpolated on a  $256^3$  grid. We use adaptive smoothing similar to SPH smoothing but with a Gaussian window truncated at  $3\sigma$ , as explained and tested in Pichon et al. (2001). This eliminates discreteness effects while keeping the best spatial resolution possible. The pixel size of the dark-matter only simulation is  $0.4 h^{-1}\text{Mpc}$ . This corresponds to  $0.47 \text{ \AA}$ , a factor 2.5 smaller than the pixel size of the FORS spectra. We have convolved the mock spectra from the numerical simulations with a Gaussian filter to match the spectral resolution of the observed spectra.

The hydrodynamical simulation is better suited for investigating the effects of thermal broadening and redshift distortions which are more relevant on small scales. This simulation has  $512^3$  dark matter particles and follows the gas dynamics on a fixed cubic grid with  $512^3$  cells. The simulation has a box-size of 40 Mpc; the mesh size is  $\sim 80 h^{-1}\text{kpc}$  which corresponds to a pixel size of  $0.07 \text{ \AA}$  at a redshift  $z = 2$ . This means that the Jeans length of the warm photo-ionized IGM is marginally resolved. We have used simulations with box-size of

---

<sup>4</sup>Note that the mean wavelength range corresponding to the Lyman- $\alpha$  forest in the observed FORS spectra,  $\sim 250 \text{ \AA}$ , corresponds to approximately twice the box-size of the simulation.

20 and 10 Mpc to check that the 40 Mpc simulation used here is not significantly affected by the fact the Jeans mass is only marginally resolved. The 40 Mpc size therefore offers the best compromise between box-size and resolution and the statistical properties of the absorption spectra studied here are sufficiently converged. The dark matter distribution is modelled with the same PM code as the dark-matter only simulation. The hydrodynamical part of the code is the same as in Chièze, Alimi & Teyssier (1998) and Teyssier, Chièze & Alimi (1998). The adiabatic hydrodynamic step is solved using directional splitting and a staggered mesh. Shock waves are approximated with the pseudoviscosity method (Von Neuman & Richmyer 1950). An additional dissipative step models the physical processes relevant for the description of gas dynamics in a photo-ionized intergalactic medium, as described in the Appendix B of Theuns et al. (1998), except that we use the heating and photoionization rates of Davé et al. (1999) which were derived from measurements by Haardt & Madau (1996). The Intergalactic Medium is highly ionized at the relevant redshifts and the dynamical evolution of the gas in the simulation depends therefore only very weakly on the amplitude of the ionizing flux characterized by its value at the Lyman limit :  $J_{21}$ . We have run the simulation with the same ionizing flux as adopted in Davé et al. (1999). However, when we are producing mock spectra, we compute the equilibrium ionic abundances in a post-processing step for a rescaled ionizing flux such as to match the observed flux distribution (see section 5.5.5). This procedure does not affect the density and temperature distribution of the gas, as specified in Theuns et al. (1998).

#### 5.5.4 The observed longitudinal and transverse correlation functions

##### Calculating correlation functions

We define the flux correlation function as in Rollinde et al. (2003) :

$$\xi_f(\theta, \Delta v) = \langle (\mathcal{F}(\theta, \lambda + \Delta\lambda) - \langle \mathcal{F} \rangle)(\mathcal{F}(0, \lambda) - \langle \mathcal{F} \rangle) \rangle_\lambda, \quad (5.7)$$

where  $\mathcal{F}$  is the normalized flux along two lines of sight with separation  $\theta$  at a mean redshift  $z$ ;  $\Delta\lambda = \lambda_0(1+z) \times \Delta v/c$ , with  $\lambda_0 = 1215.67 \text{ \AA}$  the hydrogen Lyman- $\alpha$  rest-wavelength, and  $c$  denotes the speed of light. The velocity distance corresponding to the angular separation  $\theta$  can be written as  $\Delta v_\perp = c f(z) \theta$ , where  $f(z) = c^{-1} H(z) D_A(z)$ ,  $H(z)$  is the Hubble constant at  $z$ , and  $D_A(z)$  is the angular diameter distance (see Mc Donald 2003). For  $\theta = 0$  equation (1) gives the longitudinal correlation function. In the following we will use  $H_0 = 70 \text{ km/s/Mpc}$ ,  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$  to relate scales in redshift space with angular scales. With these parameters  $\Delta v = 100 \text{ km s}^{-1}$  corresponds to  $\sim 1 \text{ arcmin}$  at  $z = 2$ .

We have excluded the wavelength range less than  $1000 \text{ km s}^{-1}$  redward of the Lyman- $\beta$  emission line when calculating the observed correlation functions to avoid contamination from the Lyman- $\beta$  forest. Likewise spectral regions less than  $3000 \text{ km s}^{-1}$  blueward of the Lyman- $\alpha$  emission line have been excluded to avoid the proximity effect (see Rollinde et al. 2005). Damped absorption

systems and metal lines which we were able to identify have been removed. The properties of identified damped absorption systems are listed in Table 5.2 and the properties of the metal lines are listed in the Appendix B published in the electronic version of the paper.

### The observed longitudinal correlation function

The observed correlation functions depend strongly on the spectral resolution of the absorption spectra unless the width of all spectral features is fully resolved (e.g Becker, Sargent & Rauch 2004).

We first consider the longitudinal correlation function obtained from the FORS data. The thick dashed line in Fig. 5.6 shows the average of the longitudinal correlation functions measured from the 58 FORS spectra in the sample. Errors are calculated from the variance of the measurements (see Section 5.5.4).

The longitudinal correlation function measured from the high resolution spectra obtained in the course of the UVES-VLT LP is shown as a thick solid curve in Fig. 5.6. The high-resolution spectra have a mean redshift of 2.39. As expected, the correlation function of the high-resolution spectra extends to higher values at small velocity separation. We also show the longitudinal correlation function obtained by Mc Donald et al. (2000) from a smaller sample of eight high-resolution spectra with a mean redshift of  $z = 2.41$  obtained with Keck-HIRES as the thin solid curve. There is excellent agreement for the two samples of high-resolution spectra at velocity separations  $\leq 300 \text{ km s}^{-1}$ . At velocity separations of  $\sim 400 \text{ km s}^{-1}$ , the longitudinal correlation function obtained from the UVES spectra appears to be larger than that obtained from the HIRES spectra. Cosmic variance or artefacts due to continuum fitting errors are two plausible explanations. Note, however, that the errors are already rather large at these separations and that the difference is probably not statistically significant.

The thin dotted curve shows the longitudinal correlation from the high-resolution VLT-UVES spectra after convolution with a Gaussian filter of  $\text{FWHM} = 220 \text{ km s}^{-1}$  to take into account the difference in resolution of the UVES and FORS spectra. It agrees very well with the FORS correlation function up to a small systematic offset.

### The observed transverse correlation function

The correlation function is calculated using the 32 pairs presented in Table 5.1. We use only spectral regions where the S/N ratio in both spectra is larger than 8. For this reason the pair J 123510.5-010746–J 123511.0-010830 is not included in the analysis as the redshift overlap of the two Lyman- $\alpha$  forests is too small to contribute in a statistically significant way. The wavelength range ( $\lambda_{\text{max}} - \lambda_{\text{min}}$ ) used to compute the transverse correlations is given in Table 5.1 together with the corresponding number of pixels. The mean flux is taken over each individual line of sight.

In Fig. 5.7 we show the observed transverse correlation function. The measurement for each quasar pair,  $\chi(\theta_i) \equiv \xi_f(\theta_i, \Delta v = 0)$ , is shown as a small solid triangle at the angular separation of the pair,  $\theta_i$ . The points with the solid er-

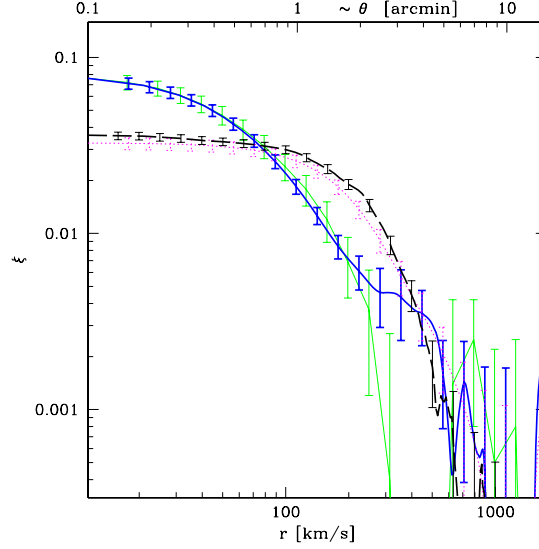


FIG. 5.6 – The observed longitudinal correlation functions derived from low resolution FORS spectra (thick dashed curve), high-resolution VLT-UVES (thick solid curve) and high resolution KECK-HIRES spectra (Mc Donald et al. 2000, thin solid curve). The thin dotted curve shows the correlation function from the VLT-UVES spectra after convolution of the spectra with a Gaussian filter of  $\text{FWHM} = 220 \text{ km s}^{-1}$  to match the resolution of the FORS spectra. Errors for the FORS and UVES spectra have been estimated from the observations as described in the text.

ror bars are a binned estimate of the transverse correlation function for which we have weighted the individual measurements with the inverse of their errors (see Section 5.5.4 for the computation of errors). Note that the first bin at the smallest separation contains only one measurement.

The transverse correlation is clearly detected on scales  $< 4 \text{ arcmin}$ . If we merge the two bins between 3 and 4 arcmin, the correlation is detected at about the  $3\sigma$  level.

Note that correlation coefficients derived from this work for some of the pairs observed by Rollinde et al. (2003) differ somewhat from what is given in their paper (see 5.1). This is due to slight differences in the reduction of the data and the determination of the continuum.

### Estimation of errors

The measurements of the longitudinal correlation reported in Fig. 5.6 are the average of the correlation function of the individual spectra  $\xi_i(\Delta v)$  (58 spectra in the case of FORS and 20 in the case of UVES),

$$\bar{\xi}(\Delta v) = \frac{1}{N} \sum_{i=1}^N \xi_i(\Delta v). \quad (5.8)$$

The error of the mean correlation function is then computed as,  $(\Delta \xi)^2 \simeq$



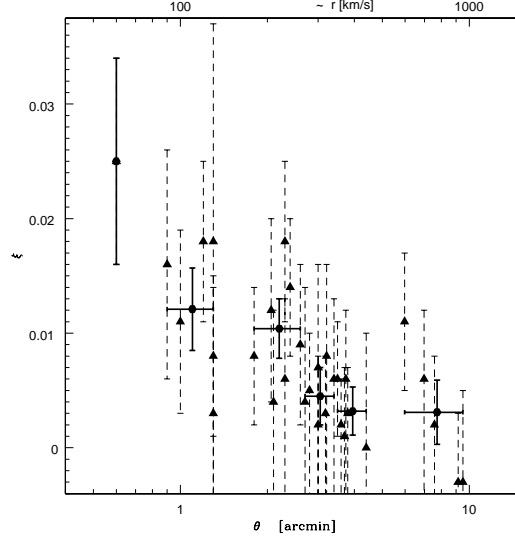


FIG. 5.7 – The observed transverse correlation coefficient for individual pairs (black triangles, dashed error bars ; see 5.1) and a binned estimate of the transverse correlation function (solid error bars). The error bars for the individual measurements are estimated using the dark-matter only simulation as explained in Section 5.5.4.

$1/N(N-1) \sum_{i=1}^N [\xi_i - \bar{\xi}]^2$ . For the transverse correlation there is only one measurement (one pair) at each separation. We therefore use the dark matter-only simulation to estimate the statistical errors. We choose samples of random sight-lines along one axis of the box carefully reproducing individual pair separations, wavelength coverage, resolution and the noise of the spectra in the observed sample. The length of the observed spectra is always larger than  $100 h^{-1}\text{Mpc}$ , and we have concatenated several lines of sight through the simulation box in order to reproduce one observed spectrum. Note that we use only one output of the simulation at  $z = 2$  and did not try to account for the moderate redshift evolution in our sample. We extract 10000 different realizations from the simulation. We then fit a Gaussian to the distribution of the values of the correlation function at each separation  $\hat{\xi}_j$  and use the rms of the distribution as estimate for the error of the transverse correlation function as reported in Table 5.1 and Figure 5.7.

Note that the errors quoted here on the flux correlation functions are only indicative because they are strongly correlated (see McDonald et al. 2000).

### 5.5.5 Comparison of observed and simulated correlation functions

#### Numerical simulations as a testbed for systematic uncertainties

The longitudinal and transverse correlation functions reflect the clustering of the underlying matter distribution in real space albeit in a somewhat in-

direct way. The correlation functions calculated from mock spectra produced from numerical simulations are an excellent tool to test the effects of resolution, redshift space distortion, thermal broadening and non-linear evolution of the gravitational clustering. We use here the  $512^3$  cell full hydro-simulation described in Section 5.5.3. When thermal broadening and redshift-distortion are taken into account, they are computed from the temperature and velocity fields of the simulation as described in Theuns et al. (1998). We produce spectra for all lines of sight along one axis of the simulation box separated by one cell. This corresponds to  $512^2$  sightlines with a length of 512 pixels each. Our estimate of the longitudinal correlation function from the simulations is obtained by averaging over these  $512^2$  individual realisations. The transverse correlation function is computed at 20 log-spaced values of  $\theta$ . We average over pairs of lines of sight for each value of  $\theta$  in the following way. For each of the  $512^2$  pixels in the y-z plane, we take the sightline parallel to the x-axis as the first spectrum of a pair. We then use two parallel lines of sight separated by a distance  $\theta$ , in the y-direction and z-direction, respectively, to compute the second spectrum to obtain two pairs of spectra. Our estimate of the transverse correlation is the average of the resulting  $2 \times 512^2$  pairs.

The normalization of the correlation function depends sensitively on the mean flux which in turn depends on the amplitude of the ionizing flux. Therefore, the mock spectra were calculated with a rescaled ionizing flux such that the probability distribution function (PDF) of the flux distribution matches that of our observed spectra at the same redshift for  $\tau = 1$ . We proceed iteratively starting with an arbitrary ionizing flux and adjusting this flux step by step till the fit is obtained. We found that this procedure is similar although more robust than the conventional procedure to match the mean flux in the Lyman- $\alpha$  forest. Indeed, the fit of the PDF minimizes the role played by overdensities and the effect of cosmic variance.

The thick and thin dotted curves in Fig. 5.8 show, respectively, the longitudinal and transverse correlation functions in real space as calculated from mock spectra produced from the full hydro-simulation. We have again used our fiducial cosmological parameters to relate velocity and angular separation. As expected the two correlation functions are almost identical. The dashed curves show the same comparison with the effect of thermal broadening included. There are now significant differences. At small scales the longitudinal correlation function exceeds the transverse correlation function while at large scales the opposite is true. The solid curves show the effect of including peculiar velocities. The corresponding redshift distortion further enhances the differences between longitudinal and transverse correlation functions. The scale dependence of the difference is similar to that due to thermal broadening but the differences are significantly larger especially at scales larger than  $200 \text{ km s}^{-1}$  or 2 arcmin.

A proper quantitative understanding of these effects with the help of numerical simulations will be essential for attempts to use the comparison of observed longitudinal and transverse correlation functions to measure cosmological parameters. This would need a full set of simulations spanning the whole range of parameters and is therefore beyond the scope of this paper.

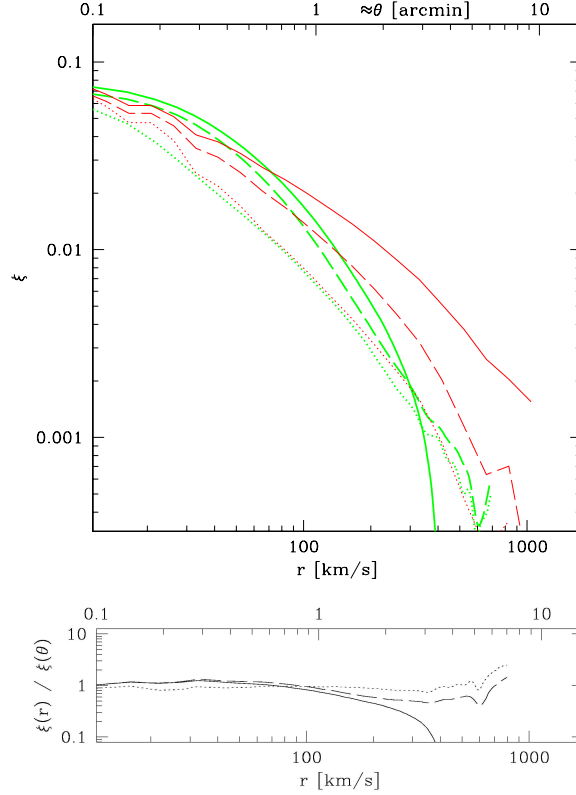


FIG. 5.8 – The longitudinal and transverse correlation functions versus velocity separation (lower x-axis) and angular separation (upper x-axis) for the full hydro-simulation at  $z = 2$ . *Upper panel* : the longitudinal correlation of the gas density (thick dotted curve) is nearly identical to the transverse correlation at the same redshift (thin dotted curve) as expected. The correlation functions differ when thermal broadening (longitudinal : thick dashed curve, transverse : thin dashed curve) and peculiar velocities (longitudinal : thick solid curve, transverse : thin solid curve) are taken into account. *Lower panel* : Ratio of the longitudinal and transverse correlation functions. Linestyles are the same as in the upper panel.

### Observed *vs* simulated correlation functions

The ability of  $\Lambda$ CDM models to reproduce the longitudinal correlation function of the Lyman- $\alpha$  forest has been demonstrated by many authors (e.g. Croft et al. 2002, Viel et al. 2002, Rollinde et al. 2003) and, as we will see below, the same is true for our simulations. The thin solid curves in Fig. 5.9 show the mean longitudinal correlation function obtained for mock spectra produced from the hydrodynamical simulation at  $z = 2$  (lower curve) and  $z = 3$  (upper curve). The curves nicely bracket the observed correlation function obtained from the high-resolution data with a median redshift of  $z = 2.39$ . Note again the slight excess of the longitudinal correlation function of the UVES data at large scales

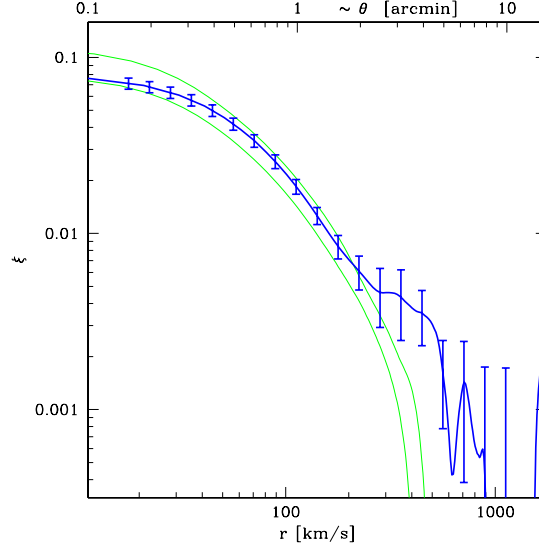


FIG. 5.9 – The observed longitudinal correlation function from the high resolution UVES spectra (thick solid curve) compared to the longitudinal correlation function as measured in the full hydro-dynamical simulation at  $z = 2$  (lower thin curve) and  $z = 3$  (upper thin curve), respectively

which is, however, probably not statistically significant.

The transverse correlation contains precious direct information on the physical size/coherence-length of the absorbing structures as it is less affected by redshift space distortions than the longitudinal correlation function. Furthermore, a comparison of the longitudinal and transverse correlation functions can – at least in principle – strongly constrain cosmological parameters in particular  $\Omega_\Lambda$ .

The thick solid curve in Figure 5.10 shows our estimate of the transverse correlation function from the full hydro-dynamical simulation at  $z = 2$ . It agrees well with our measurement of the observed transverse correlation function (mean redshift  $\sim 2.1$ ) which is shown as the solid dots with error bars. The thin dashed curve shows the prediction of linear theory (Kaiser 1987; Mc Donald & Miralda-Escudé 1999) for the cosmological parameters assumed for the hydro-simulation. The thin solid and dotted curves show the prediction of linear theory for  $\Omega_m = 0.1$  and  $\Omega_m = 1.0$ , respectively (assuming a flat universe :  $\Omega_\Lambda + \Omega_m = 1$  and adjusting other parameters to fit the data). The linear theory predictions are normalized so that the longitudinal correlation function is best fitted for  $\Delta v > 200 \text{ km s}^{-1}$ . As expected, the linear predictions agree reasonably well with the numerical simulation at large scales but underpredicts the correlation function substantially at small scales. The non-linear effects of gravitational clustering are clearly visible in the observed transverse correlation function.

Despite the larger sample (about three times more pairs at  $\theta < 3 \text{ arcmin}$  than in Rollinde et al. 2003) and the correspondingly smaller errors, we cannot yet distinguish between different values of  $\Omega_m$ . This confirms the predictions by Rollinde et al. (2003) and Mc Donald (2003) that significant constraints on  $\Omega_\Lambda$

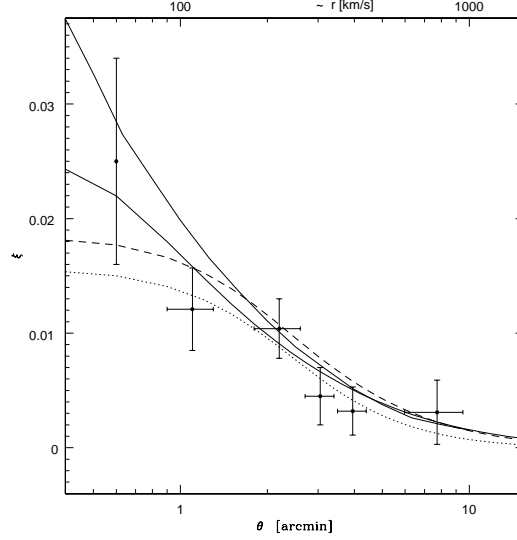


FIG. 5.10 – The binned estimate of the observed transverse correlation function (solid error bars) is shown together with the estimate of the transverse correlation function from the full hydro-simulation (thick solid curve) and linear predictions for  $\Omega_m = 0.1, 0.3$  and  $1$  (thin solid, dashed and dotted lines respectively, assuming a flat universe :  $\Omega_\Lambda + \Omega_m = 1$  ; see text). The normalization of the flux in the simulation is fixed in order to reproduce the observed flux PDF (Section 5.5.5). The linear theory predictions are normalized to reproduce the longitudinal correlation function at large scales.

require a larger number of pairs. Using the (cross) power spectrum instead of the transverse and longitudinal correlation functions, Mc Donald (2003) estimates that of the order of  $13(\theta/1')^2$  quasar pairs on scales up to 10 arcmin are necessary to perform the test. In addition, performing the Alcock & Paczyński test using the correlation functions at small scales ( $\leq 3$  arcmin or  $300 \text{ km s}^{-1}$  at  $z = 2$ ) will require the use of a large suite of full hydro-dynamical simulations.

### 5.5.6 Metal absorption systems

#### Identifying metal lines

Our sample is also well suited to study the spatial distribution of the gas responsible for the associated metal absorption in QSO spectra. We have manually identified and fitted metal lines in all spectra. The corresponding line lists are compiled in Appendix B (published in the electronic version of the paper) where we report the observed equivalent width, the pixel signal-to-noise ratio at the position of the absorption, the observed wavelength and corresponding

QSO name	$z_{\text{em}}$	$\theta$	SNR	$\lambda$	Number of pixels	$\xi(0, \theta)$	$\sigma$
*Q2139-4504B	3.255	0.600	24.4	4381- 4878	420	0.025	0.009
Q2139-4504A	3.055		15.2				
J 123510.5-010746	2.785	0.740	51.4	3897- 3904	7	-	-
J 123511.0-010830	2.235		26.2				
J 232800.7-271655	2.378	0.900	14.3	3756- 4038	240	0.016	0.010
J 232804.4-271713	2.357		14.3				
UM680	2.144	1.000	50.0	3235- 3699	9641 <sup>a</sup>	0.011	0.008
UM681	2.122		50.0				
J 031036.4-305108	2.552	1.200	20.9	3658- 4249	501	0.018	0.007
J 031041.0-305027	2.532		25.4				
J 135001.7-011703	2.657	1.300	15.5	3765- 3821	48	0.018	0.019
J 135003.0-011819	2.177		18.5				
J 000852.7-290044	2.699	1.300	64.1	3808- 4323	317	0.008	0.007
J 000857.7-290126	2.593		34.2				
J 214507.0-303046	2.532	1.300	37.5	3636- 3869	177	0.003	0.011
J 214501.6-303121	2.216		27.3				
J 005852.4-272933	2.565	1.800	23.9	3671- 4284	493	0.008	0.006
J 005859.1-273038	2.561		25.5				
*Q0103-294B	2.190	2.063	25.5	3602- 3827	191	0.012	0.008
Q0103-294A	2.182		28.3				
*Q2129-4653B	2.222	2.100	22.4	3603- 3856	215	0.004	0.008
Q2129-4653A	2.206		16.8				
J 031054.7-293436	2.281	2.300	18.3	3601- 3833	197	0.006	0.007
J 031103.0-293306	2.187		29.3				
J 102827.1-013641	2.393	2.300	13.4	3728- 3954	193	0.018	0.007
J 102832.6-013448	2.287		17.0				
J 111201.8-013018	2.549	2.400	48.7	3654- 3954	246	0.014	0.006
J 111200.4-013242	2.292		28.1				
*Q0236-2411	2.260	2.600	20.3	3602- 3860	219	0.009	0.007
Q0236-2413	2.211		19.4				
J 125556.9+001848	2.108	2.700	19.9	3602- 3708	90	0.004	0.010
J 125606.3+001728	2.083		20.0				
J 013734.2-303802	2.481	2.800	24.4	3602- 4005	341	0.005	0.005
J 013734.2-304050	2.329		29.2				
J120725.9-024519	2.676	3.000	33.2	3785- 3904	101	0.007	0.009
J120734.5-024725	2.245		20.0				
J 095810.9-002733	2.559	3.000	21.4	3664- 4047	287	0.002	0.006
J 095800.2-002858	2.364		14.2				
J 223850.1-295612	2.448	3.180	32.0	3602- 4062	390	0.003	0.005
J 223850.9-295301	2.377		34.2				
J 141124.6-022943	2.710	3.210	43.2	3820- 3971	128	0.008	0.008
J 141117.3-023222	2.301		35.3				
J 023836.9-282310	2.565	3.400	31.7	3677- 3899	175	0.006	0.007
J 023849.0-282101	2.242		56.7				
Q 1207-1057	2.450	3.500	33.8	3601- 3975	318	0.006	0.005
Q 1206-1056	2.305		24.7				
J 215225.8-283058	2.741	3.600	38.0	3851- 4494	545	0.002	0.004
J 215240.0-283251	2.736		25.3				
J 112116.1+003112	2.205	3.700	19.0	3603- 3834	197	0.001	0.006
J 112108.2+003420	2.188		22.8				
J144245.7-023906	2.551	3.740	24.3	3656- 4011	232	0.006	0.006
J144245.6-024251	2.334		21.2				
J 230318.4-290120	2.587	3.800	35.5	3693- 4285	500	0.003	0.004
J 230301.6-290027	2.562		26.9				
*FOCAP QSF :01	2.267	4.400	21.5	3602- 3673	61	0.000	0.010
FOCAP QSF :04	2.054		19.0				
*Q0102-2931	2.212	5.974	18.5	3603- 3837	199	0.011	0.006
Q0103-294B	2.190		25.5				
*Q0102-2931	2.212	6.977	18.5	3602- 3827	191	0.006	0.006
Q0103-294A	2.182		28.3				
*Q0102-293	2.441	7.585	24.8	3602- 3827	191	0.002	0.006
Q0103-294A	2.182		28.3				
*Q0102-293	2.441	9.152	24.8	3603- 3837	199	-0.003	0.006
Q0103-294B	2.190		25.5				
*Q0102-293	2.441	9.506	24.8	3602- 3864	222	-0.003	0.008
Q0102-2931	2.212		18.5				

TAB. 5.1 – Properties of the 33 QSO pairs : QSO names, emission redshifts, angular separation on the sky, mean S/N ratio over the wavelength range of interest, wavelength range (in Å) over which the correlation is calculated, corresponding number of pixels, value of the correlation function and errors (see Section 4.4). (a) UVES data (D’Odorico et al. 2002) with 0.04 Å per pixel instead of 1.18 Å for FORS data. (\*) QSO pair observed by Rollinde et al. (2003).

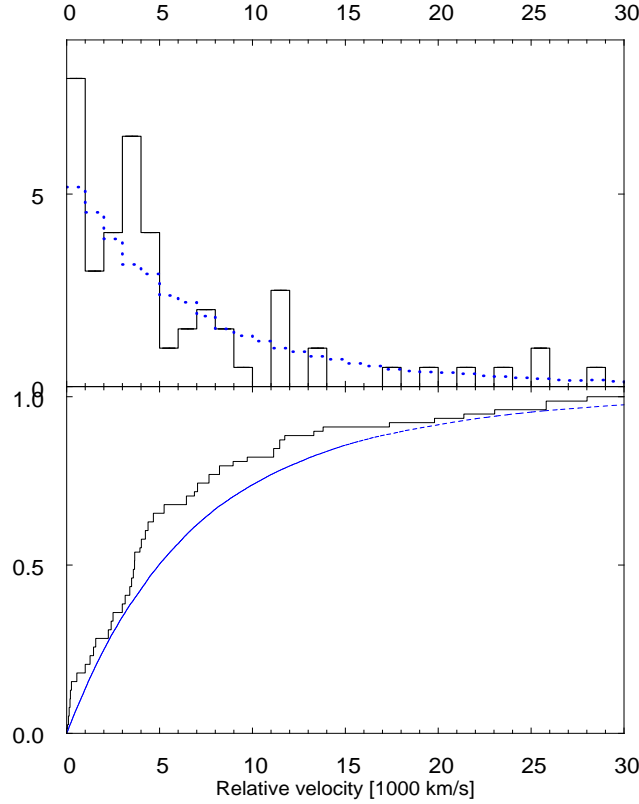


FIG. 5.11 – *Upper panel* : Histogram of the velocity separations between two closest-neighbor C IV absorption systems in our sample of QSO pairs. The distribution expected from a randomly distributed population of C IV systems is shown as the dashed curve. *Lower panel* : the cumulative distributions of the observed sample of C IV systems (thick curve) is compared to the cumulative distribution of a randomly distributed population of C IV systems (dashed curve). There is a 18% chance probability (KS test) that the distributions differ that much if the two sample are drawn from the same population.

redshift for each of the identified metal lines. If applicable the upper limit on the equivalent width of a possible absorption at the same position in the spectrum of the second quasar of the pair is also given.

### The correlation of C IV systems along adjacent lines of sight

We consider only absorption lines with rest-frame equivalent width  $W_r > 0.1 \text{ \AA}$  and redshift intervals common to both lines of sight of a QSO pair. We do not consider systems where N V is detected as these systems are most probably associated with the quasar (see e.g. Petitjean, Rauch & Carswell 1994). We then select the lines that are located at more than  $3000 \text{ km s}^{-1}$  blueward of the QSO C IV emission line and at more than  $1000 \text{ km s}^{-1}$  redward of the Lyman- $\alpha$  emission line. We end up with a sample of 139 C IV systems for a redshift path  $\delta z = 38$ , corresponding to a density of 3.7 systems per unit redshift. We apply the Nearest-Neighbor method, as described in Young et al. (2001) and

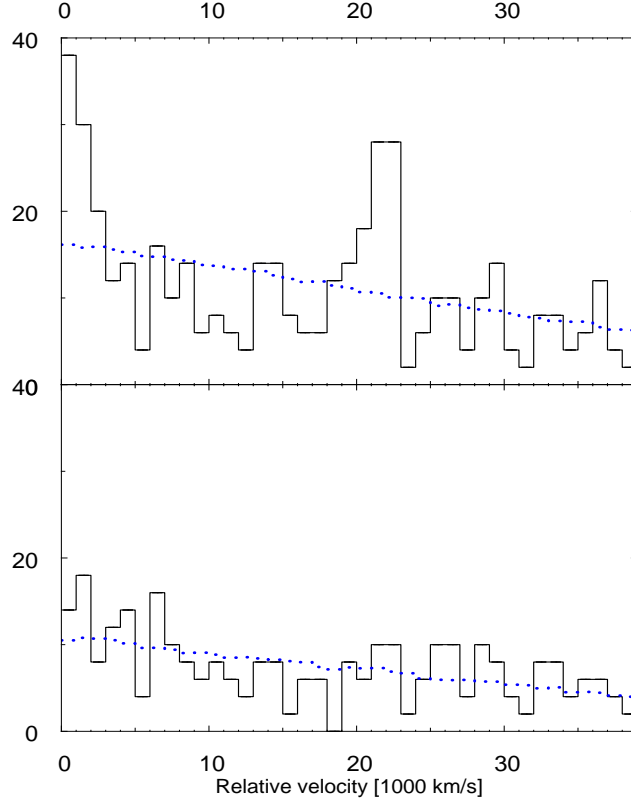


FIG. 5.12 – Observed Longitudinal correlation function of C IV systems computed along the 58 lines of sight. The distribution expected from a population of randomly distributed C IV systems is shown as the dashed curve. *Upper panel* : the whole sample is used. There is an excess on small scales and around  $20000 \text{ km s}^{-1}$ . This is due to an excess of correlation along the lines of sight to Q 0103–294A,B; Q 0102–2931 and Q 0102–293 located in the same field (group of quasars). *Lower panel* : same but without the group.

Aracil et al. (2002) to the corresponding list of C IV systems. For each absorption line along one QSO line of sight, we search the adjacent QSO line of sight for the nearest (in velocity) absorption line and construct the histogram of the corresponding velocity differences (see Fig. 5.11). Our complete sample contains 25 and 39 associations with velocity separations smaller than  $5000$  and  $30000 \text{ km s}^{-1}$ , respectively. To estimate the possible excess of correlation with respect to randomly located absorption lines, we produced 10000 simulated line lists drawn from a population of randomly redshifted lines, taking the same number of lines and the same wavelength ranges as in the observed spectra. The results of applying the same method to the simulated line lists are given as dotted lines in Figure 5.11.

In the lower panel of Fig. 5.11 we compare the cumulative distributions of the velocity differences from our observed sample and from a randomly located population of C IV absorbers. A KS test gives a 18% chance probability that the difference between the two distributions is larger than what is observed if



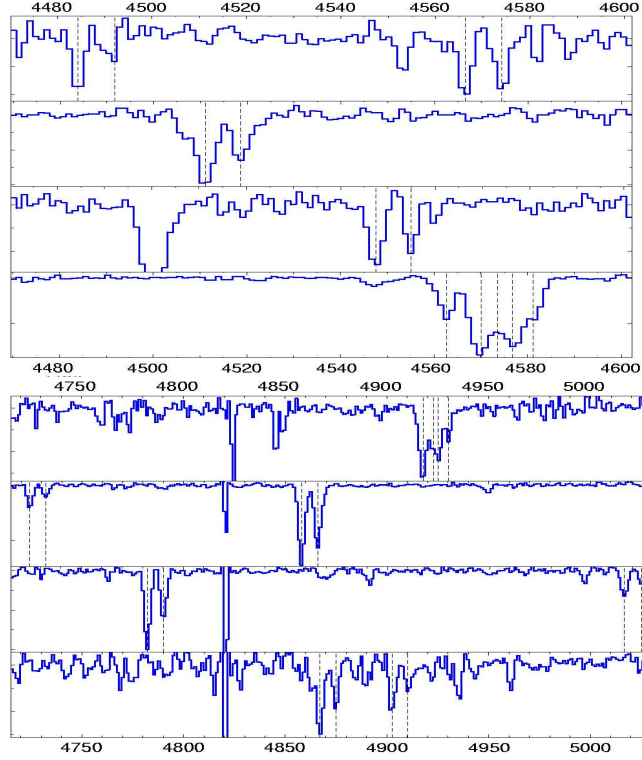


FIG. 5.13 – Two sections of the spectra of Q 0103–294A&B (separated by 1.3 arcmin) and Q 0102–2931 and Q 0102–293 (separated by about 6.5 arcmin from Q0103–294A,B, see Table 5.1) in that order from top to bottom. The mean redshifts are 1.93 and 2.14 in the upper and lower panels, respectively. The positions of C IV systems are indicated by dashed vertical lines. There is an overdensity of C IV systems along these lines of sight over the redshift range  $1.5 \leq z \leq 2.2$  and on a spatial scale larger than 10 arcmin.

the two samples are drawn from the same population. There is a possible small excess of clustering of C IV systems on scales smaller than  $5000 \text{ km s}^{-1}$ . This scale is larger than the typical correlation length, of about  $1000 \text{ km s}^{-1}$ , seen in the longitudinal correlation function of C IV systems (Rauch et al. 1996; Pichon et al. 2003; Boksenberg, Sargent & Rauch 2003; Scannapieco et al. 2006). This is due to an excess of associations in the bin  $\Delta v \sim 4000 \text{ km s}^{-1}$ . The corresponding C IV associations are located in the peculiar field containing the quartet Q 0103–294A&B, Q 0102–2931 and Q 0102–293 (all four quasars are separated by less than 10 arcmin). In fact, 9 out of the 25 associations with  $\Delta v < 5000 \text{ km s}^{-1}$  (see next Section) are found in front of these quasars. If we remove this quartet from the sample, the KS probability is increased to 25%. To ascertain the overdensity in this field, we show in Fig. 5.12 the longitudinal correlation function for the whole sample (upper panel) and for the sample without the group (lower panel). There is indeed a strong excess in the correlation function of the whole sample for  $\Delta v < 3000 \text{ km s}^{-1}$  and around  $20000 \text{ km s}^{-1}$  which disappears when the group is removed from the sample

The pairs in the current sample have a mean separation larger than 2 arcmin (see Table 5.1). The correlation at smaller separations can be expected to be larger. Indeed, comparing the results of  $N$ -body simulations to high spectral resolution observations, Scannapieco et al. (2006) have shown that, at  $z \sim 3$ , the longitudinal C IV correlation function is consistent with a model where C IV is confined within bubbles of typical radius  $\sim 2$  Mpc comoving surrounding halos of mass  $\sim 10^{12} M_{\odot}$ . At this redshift, this corresponds to a separation of about  $\sim 2$  arcmin. Unfortunately, the small size of our sample prevents any attempt to consider only small separations. Our result is, however, consistent with their findings.

### Peculiarities

It is interesting to note that there is an overdensity of C IV pairs in front of the quartet Q 0103–294A&B, Q 0102–2931 and Q 0102–293. The density of C IV systems along the four lines of sight (6.4 per unit redshift) and the number of coincidences within  $4000 \text{ km s}^{-1}$  is about twice larger than the mean density of coincidences in the overall sample. In Fig. 5.13 we plot two portions of the spectra of Q0103–294A,B (separated by 1.3 arcmin) and Q 0102–2931 and Q 0102–293 (separated by about 6.5 arcmin from Q0103–294A,B, see Table 5.1). The wavelength ranges of the two portions are centered on  $z = 1.92$  and  $2.14$ , respectively. Note, however, that the overdensity of C IV systems along these lines of sight extends over the much larger redshift range  $1.5 \leq z \leq 2.2$  and over a spatial scale larger than 10 arcmin (see Appendix B). There are two more peculiarities occurring along these lines of sight. There are no C IV systems between  $z = 1.955$  and  $2.051$  and there is a quasi-spherical structure of reduced H I absorption with radius  $\sim 12.5h^{-1}$  Mpc at  $z \sim 1.992$  in front of the quartet (Rollinde et al. 2003). Note also that the correlation between the Lyman- $\alpha$  forests of Q 0102–2931 and Q 0103–294B, two quasars of the group with  $\sim 6$  arcmin separation, is measured to be quite high ( $\xi = 0.11$ , see Table 5.1).

A similar overdensity of C IV systems has been observed in the field of Tol 1037–2704 (e.g. Jakobsen et al. 1986; Dinshaw & Impey 1996; Lespine & Petitjean 1997). The overdensity of C IV systems in this field extends over the redshift range  $\sim 1.5$ – $2.2$  and a transverse scale  $> 15$  arcmin and has been interpreted as being due to the presence of a supercluster. The dimensions of this supercluster would be at least 80 and  $30h^{-1}$  Mpc along and perpendicular to the line of sight, respectively. To our knowledge no deep imaging of this field exists. Another overdensity of C IV systems has been reported in front of PKS 0237–233 (Sargent, Boksenberg & Steidel 1988; Foltz et al. 1993). The overdensity reported in this work in the field around Q 0103–294A,B may give new clues to solve the puzzle of the origin of these overdensities extending over very large scales as the four quasars constituting the quartet are very close to each other. Deep infra-red imaging should be performed in the field to search for any concentration of objects in the corresponding redshift range. High spectral resolution observations of the quasars would allow a more detailed investigation of the nature of these C IV systems.

he pair J 000852.7-290044/J000857.7-290126 is also peculiar as the two lines of sight show 9 and 4 C IV systems respectively, corresponding to 3.5 and 1.5 times the mean density of systems. J 000852.7-290044 shows a BAL systems and it would be interesting to question the intervening origin of some of the narrow systems (Srianand & Petitjean 2001).

### 5.5.7 Conclusions

We have obtained VLT-FORS observations of a large sample of 32 pairs of QSO with separations in the range  $0.6 < \theta < 10$  arcmin building on the smaller sample of Rollinde et al. (2003) of 11 pairs. We present measurements of the transverse and longitudinal correlation functions from this sample. We further use a large box-size DM only simulation and a somewhat smaller full hydrodynamical simulation to investigate the effect of spectral resolution, thermal broadening and peculiar motions on the correlation function and to determine realistic error estimates. The longitudinal correlation function from the FORS sample is in good agreement with that obtained from UVES high-resolution data if the effect of the different spectral resolutions is taken into account.

The transverse correlation is detected at the  $3\sigma$  level up to separations of about  $\sim 3-5$  arcmin. The sample is sufficiently large to obtain a binned estimate of the average correlation function which has about a factor 1.5-2 smaller errors than the smaller sub-sample described in Rollinde et al. (2003). The shape and correlation length of the transverse correlation function of the absorbing gas is in good agreement with expectations for absorption by density fluctuations in the warm photo-ionized Intergalactic Medium as described in CDM-like structure formation models. Our measurement of the transverse correlation function is thus an important further independent confirmation that the Lyman- $\alpha$  forest is indeed caused by the filamentary and sheet-like structures of the cosmic web predicted by these models.

We then use the numerical simulations and predictions of linear theory to assess prospects of using the transverse correlation function for a variant of the Alcock & Paczyński test to determine cosmological parameters. In agreement with predictions of previous theoretical studies we find that our sample is still too small for this purpose. The improved errors of our larger sample compared to the sub-sample of Rollinde et al. (2003) suggest however that meaningful constraints on  $\Omega_\Lambda$  can be obtained. For this, a larger sample and a careful analysis of the systematic uncertainties with a large suite of full hydrodynamical simulations are necessary. Mc Donald (2003) estimated that this requires a sample of  $13(\theta/1')^2$  pairs on scales up to 10 arcmin.

We have also used our sample to investigate the transverse and longitudinal correlation functions of C IV absorption systems on the scales probed by our pairs, but did not detect any signal. This is not surprising as most of the separations are larger than 2 comoving Mpc. This is larger than is expected for the size of metal-enriched bubbles surrounding massive haloes (e.g. Scannapieco et al. 2006). We have, however, detected a prominent overdensity of C IV systems in front of the quartet Q 0103-294A&B, Q 0102-2931 and Q 0102-293 which extends over the redshift range  $1.5 \leq z \leq 2.2$  and over a spatial scale larger than

QSO	$w_{\text{obs}}$ ( $\text{\AA}$ )	$\lambda_{\text{obs}}$ ( $\text{\AA}$ )	$z_{\text{abs}}$
J 135003.0–011703	29.6	4055.77	2.336
J 144245.6–024251	21.4	3911.52	2.218
J 000852.7–290044	22.2	3955.77	2.254
J 000857.7–290126	20.0 :	4243.08	2.490
Q 2139–4504B	55.3	5071.54	3.172

TAB. 5.2 – Damped Lyman- $\alpha$  system candidates detected in the survey

10 arcmin. This suggests the presence of a high-redshift cluster in this field and makes it a prime target for deep infra-red imaging.

### Acknowledgements

We thank D. Weinberg for useful discussions and for providing the numerical tables of heating and photoionization rates used in the hydrodynamic simulation presented in this paper. We thank an anonymous referee for a thorough reading of the manuscript. The simulations were performed as part of a Numerical Investigations in Cosmology group task in the framework of the HORIZON project. Computer time for the simulations was allocated by the scientific council of IDRIS, Orsay. FC thanks IUCAA-Pune (India) for hospitality during the time part of this work has been completed and ESO-Vitacura for a PhD studentship.



# Conclusion

*"There's never enough time to do all the nothing you want."*  
– Calvin

L'étude du milieu intergalactique grâce aux absorptions laissées dans les spectres des quasars à grand décalage spectral ne peut aujourd'hui évoluer qu'en empruntant deux chemins : tout d'abord on peut collecter un nombre encore plus grand de lignes de visée afin d'augmenter les statistiques, ce qui correspond à ce qui a été fait avec le SDSS. Celui-ci a permis d'augmenter de façon très significative les lignes de visées et ainsi de produire des statistiques sur différents types de systèmes d'absorption, par exemple en étudiant l'évolution de la forêt Lyman- $\alpha$  (e.g. McDonald et al. 2002) ou bien en sélectionnant les systèmes lorentziens (e.g. Prochaska et al. 2005). Néanmoins, les grands relevés couvrant une grande partie du ciel sont limités en magnitude et c'est là qu'intervient la seconde voie : l'étude plus profonde de champs particuliers afin d'obtenir une grande densité de lignes de visée. Avec cette densité particulière d'objets il est alors possible d'étudier les corrélations d'objets dans les champs observés. Ces deux aspects nécessitent une recherche exhaustive des quasars. C'est en ce sens que la recherche des quasars représente la première pierre d'un édifice qui reste encore à construire, et cette thèse s'est exercée à sélectionner et à mettre en place les outils nécessaires à une recherche efficace dans le CFHTLS.

Dans un premier temps j'ai comparé les diagrammes couleur-couleur et les méthodes d'adaptation de DSE afin de déterminer lequel des deux permettrait de tirer pleinement profit des qualités photométriques du relevé, le CFHTLS, dans lequel nous allons effectuer notre recherche. La description des outils englobe l'ensemble des possibilités de sélections offertes par le relevé. Tout d'abord j'ai présenté les applications actuelles des diagrammes couleur-couleur essentiellement développées par l'équipe du SDSS. J'ai détaillé le fonctionnement de leur algorithme en montrant les limites de la sélection, à savoir la limite en magnitude de la complétude à  $i < 19.1$  et sa spécificité aux couleurs du SDSS. Les performances de leur sélection atteignent une efficacité de 90%. Ensuite est venue la présentation du choix final de l'outil, la méthode d'adaptation à des DSE, en détaillant son application sur certains relevés, comme COMBO-17.

J'ai brièvement décrit le mécanisme de fonctionnement d'HyperZ avec sa mise en application aux données issues d'un relevé test : le SDSS. J'ai conduit sur celui-ci plusieurs tests permettant de mettre en place notre sélection. Le chapitre 4 présente l'ensemble des tests effectués sur le SDSS ainsi que les choix qu'ils m'ont amené à faire. Les résultats de ces tests sur notre méthode de

sélection ont permis de calibrer les outils utilisés. Cette calibration permet de retrouver plus de 95% des quasars présents dans le SDSS ( $i < 19.1$ ), avec une efficacité supérieure à 50%. Cette sélection s'accompagne de la détermination des décalages spectraux photométriques des quasars sélectionnés.

J'ai montré comment fonctionnait l'application de notre sélection aux données issues du CFHTLS, selon les différentes magnitudes auxquelles nous pourrions nous limiter pour différents suivis spectroscopiques.

Afin de pouvoir estimer la qualité de la sélection, il faudra attendre un suivi spectroscopique complet des candidats quasars. Néanmoins les dernières données, notamment issues du VVDS, permettent de se rendre compte que la plupart des quasars sont retrouvés par notre sélection, ce qui nous renseigne au moins sur la complétude, de plus de 90%, de la sélection.

Le dernier chapitre représente la base du travail qu'il faudra effectuer une fois que l'on aura collecté suffisamment de lignes de visée. En effet, nous avons montré qu'avec un nombre faible de paires de quasars, il était possible d'accorder observations et simulations. La corrélation transverse trouvée à partir de notre échantillon de 32 paires de quasars est parfaitement cohérente avec les simulations et la corrélation est détectée à  $3\sigma$  pour des séparations de l'ordre de  $\sim 3-5$  arcmin.

La forme et la longueur de corrélation de la fonction transverse du gaz responsable des absorptions est en accord avec les modèles d'absorption par des fluctuations de densité dans le milieu intergalactique chaud, comme il est décrit dans les modèles de matière noire froide de formation des structures. Notre mesure de la fonction de corrélation transverse est de plus une autre confirmation importante que la forêt Lyman- $\alpha$  est en effet la conséquence de la structure filamentaire de la toile cosmique prédite par ces modèles.

De plus la recherche des métaux et l'étude des structures qui les renferment sera enrichie par la collecte des lignes de visée proches, des groupes de quasars. La fonction de corrélation transverse peut être aussi calculée sur des échantillons d'absorbants métalliques afin de nous renseigner sur la distribution des métaux et sur les structures qui les abritent. J'ai effectué la recherche des absorbants métalliques dans notre échantillon de paires de quasars (cf. Annexe C). Le chapitre 5 présente les résultats pour le calcul de la fonction de corrélation transverse en utilisant les systèmes C IV détectés et sélectionnés dans notre échantillon. L'échantillon de paires de quasars utilisées a une séparation moyenne supérieure à 2 arcmin. On s'attend à ce que la corrélation à plus petite séparation soit plus grande. Scannapieco et al. (2006) proposent un modèle à un décalage spectral  $z \sim 3$  où les systèmes C IV sont confinés dans des bulles d'un rayon typique de  $\sim 2$  Mpc comobile ce qui correspond à ce décalage spectral à une séparation de l'ordre de  $\sim 2$  arcmin. Mais la faible taille de notre échantillon ne permet pas de sonder les faibles séparations; néanmoins nos résultats sont cohérents avec leurs conclusions.

Ce travail sur les paires de quasars nous a aussi permis de mettre en exergue la voie qui a été emprunté jusqu'ici par les différentes études : il s'agit de l'examen de ligne de visée particulière. Nous avons détecté la présence d'une surdensité de systèmes C IV en face du groupe Q 0103-294A&B, Q 0102-2931 et Q 0102-293 qui s'étend sur un domaine en décalage spectral de  $1.5 \leq z \leq 2.2$  et

sur une échelle plus grande que 10 arcmin. Ceci suggère la présence d'un amas à grand décalage spectral dans ce champ.

Dans l'état actuel des connaissances, la recherche de quasars est désormais un outil bien maîtrisé, qui peut s'adapter aux différents relevés sur laquelle elle va être menée. La recherche des groupes de quasars pour l'étude du milieu intergalactique en est à ses prémices ; les bases sont posées et les moyens observationnels existent. Il faut que ceux-ci s'accompagnent de plusieurs avancées :

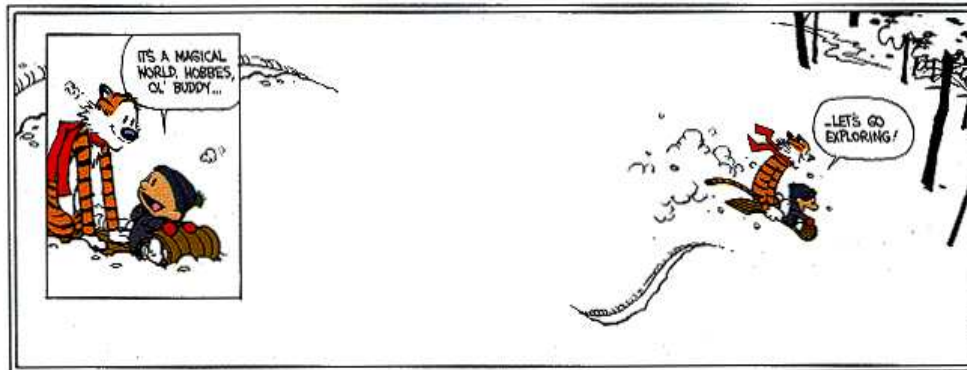
- La création d'une base de données de spectres de quasars de grande qualité. La recherche effectuée dans cette thèse en constitue la première étape. Les relevés systématiques en spectroscopie, qui restent encore trop coûteux pour être menés sur de très grandes régions du ciel, représenteront dans l'avenir la source des nouvelles bases de données.

- Une méthode de réduction homogène pour ces données. La maîtrise des erreurs systématiques nécessite que tout l'échantillon ait subi la même réduction. Dans le cas des paires, pour l'estimation du continu des quasars, nous avons eu recours à un procédé semi-automatique et nous avons pu comparer notre échantillon à celui du 'Large Program'

- Des méthodes d'analyses reproductibles et fiables, en passant par la comparaison à des simulations. Par exemple l'analyse faite sur l'échantillon constitué de 32 paires a été éprouvée et les simulations ont permis une estimation des erreurs sur la fonction de corrélation transverse.

Au final, l'arrivée de données de plus en plus abondantes nécessite des outils qui ont été calibrés sur de plus petits échantillons. Les stratégies d'observation vont dépendre des traitements disponibles en aval, afin de sélectionner les objets étudiés et de les analyser de façon homogène. C'est sous ces conditions que le projet de cartographier l'Univers en trois dimensions à de grands décalages spectraux pourra être mené à bien.





# Annexe A

## Résumé

Grâce à la très grande luminosité des quasars, ceux-ci sont utilisés par les cosmologistes comme des lampes dont les spectres permettent, via l'étude des raies d'absorption, d'extraire des informations sur l'état de la matière située le long de la ligne de visée. Il est envisageable de pousser cette recherche d'information à la reconstruction en trois dimensions du milieu intergalactique, en ayant recours à des groupes de quasars, permettant ainsi l'utilisation d'une information transverse. Pour pouvoir utiliser ces groupes encore faut-il se donner la peine de les chercher.

Le principe même de l'identification d'objets, les quasars pour ce qui nous concerne, est de comparer une base de connaissances à de nouvelles observations. Les bases de connaissances sont l'ensemble des spectres de quasars collectés jusqu'à aujourd'hui (pour l'essentiel provenant du SDSS, Sloan Digital Sky Survey); les nouvelles observations qui sont utilisées dans cette thèse proviennent du relevé CFHTLS (Canada France Hawai Telescope Legacy Survey).

Les méthodes de sélection des quasars ont beaucoup évolué ces dernières décennies. La méthode d'adaptation de distribution spectrale d'énergie (DSE) semble plus adaptable et plus fiable d'un relevé à un autre; elle se montre moins spécifique au relevé en présentant des résultats similaires.

Les tests menés sur le SDSS avec HyperZ permettent d'estimer dans un premier temps l'efficacité de notre sélection. La base de test fait 24 degrés carrés, et a été choisie pour avoir la plus grande partie de ses objets suivis en spectroscopie. Le catalogue photométrique contient  $\sim 35500$  objets à la magnitude limite  $i < 19.1$  et notre sélection retrouve une grande partie (95%) des quasars. Les tests menés sur les bases du SDSS nous permettent de disposer de plusieurs types de sélection selon les stratégies des suivis spectroscopiques. Pour chacune d'entre elles nous avons une estimation de l'efficacité. Je présente comment ces critères de sélection peuvent être appliqués au relevé CFHTLS et quelles en sont les contraintes. Sur certains champs profonds du CFHTLS nous disposons de petites bases de données de quasars qui nous permettent de quantifier dans une certaine mesure l'efficacité de notre méthode. Au final la complétude de notre sélection est estimée à plus de 90% sur le relevé CFHTLS.

Le dernier chapitre représente la base du travail qu'il faudra effectuer une fois que l'on aura collecté suffisamment de lignes de visée. En effet nous avons

montré qu'avec un nombre faible de paires de quasars, il était possible d'accorder observations et simulations, ainsi que d'obtenir un échantillon suffisant pour l'estimation de la corrélation longitudinale. La corrélation transverse trouvée à partir de notre échantillon de 32 paires de quasars est parfaitement cohérente avec les simulations et la corrélation est détectée à  $3\sigma$  pour des séparations de l'ordre de  $\sim 3-5$  arcmin.

**mots-clés** : quasars - distribution spectrale d'énergie - milieu intergalactique - SDSS - CFHTLS - systèmes d'absorption - forêt Lyman-alpha

# Annexe B

## Abstract

As quasars are among the brightest sources in the sky, they are used by cosmologists like lanterns. Their spectra allow them to extract, through the study of absorption lines, information on the equation of state of the matter along the line of sight. We can go further and reconstruct the 3D structure of the intergalactic medium by using groups of quasars, providing thus information on the transverse direction. The first step is then to look for such groups.

To identify objects, especially quasars, by their colors, one compares a database of similar objects to new observations. The databases consist of all the spectra of quasars collected through the years (mainly coming from the SDSS, Sloan Digital Sky Survey). The new observations used in this thesis come from the CFHTLS (Canada France Hawa Telescope Legacy Survey).

Methods dedicated to the selection of quasars have undergone a lot of evolution this last decades. One of them is especially devoted to our purpose by its reliability and adaptability : fitting to Spectral Energy Distribution (SED).

We use the SDSS as a test-bench for our selection. These tests allow us to estimate the efficiency of our selection. We use a field of 24 square degrees from the SDSS, which contains  $\sim 35000$  objects down to a magnitude of  $i < 19.2$ . Our selection recovers most (95%) of the SDSS quasars in this field. These tests provide also different kinds of selection according to the final follow-up strategy that can be applied, and for each of these selections I give an estimate of our efficiency. I show how to apply these selection criteria on the CFHTLS. For small parts of the CFHTLS we have a database of quasars which allows us to quantify the efficiency of our methods in the CFTHLS. We estimate the completeness of our selection at 90% on this survey.

The last chapter describes the kind of work that will be needed once we will have collected enough line of sight. We have shown with our sample of 32 pairs of quasar that the shape and correlation length of the transverse correlation function of the absorbing gas is in good agreement with expectations for absorption by density fluctuations in the warm photo-ionized Intergalactic Medium as described in CDM-like structure formation models. In particular, the transverse correlation is detected at the  $3\sigma$  level up to separations of about  $\sim 3-5$  arcmin.

**keywords :**

quasars - spectral energy distribution - SDSS - CFHTLS - intergalactic medium  
- absorption lines - Lyman- $\alpha$  forest

## Annexe C

# Annexes de l'article

### C.1 Comments on individual lines of sight

In this Section, we comment on peculiarities of individual observed lines of sight. The quasar emission redshifts are determined by fitting a Gaussian profile to the C IV emission line when present in the spectrum or to the Lyman- $\alpha$  emission line otherwise. Damped Lyman- $\alpha$  systems are listed in Table 5.2 and identified metal lines are given in Tables gathered in the Appendix.

#### C.1.1 J 000852.7-290044–J 000857.7-290126

QSO J 000852.7-290044 has a BAL system close to the emission redshift with two strong components seen in N V, O VI and C IV. In addition there is a damped Lyman- $\alpha$  system at  $z_{\text{abs}} = 2.254$  ( $W_{\text{obs}} = 22.2 \text{ \AA}$ ) with strong associated metallic absorption. There is no corresponding absorption toward J 000857.7-290126 but another damped Lyman- $\alpha$  system is detected at  $z_{\text{abs}} = 2.490$  ( $W_{\text{obs}} = 20.0 \text{ \AA}$ ) with O VI and O I associated absorption. C IV absorption is detected at  $z_{\text{abs}} = 2.218$  toward J 000852.7-290044 and at  $z_{\text{abs}} = 2.215$  toward J 000857.7-290126 (i.e. with a velocity difference of only  $294 \text{ km s}^{-1}$ ).

#### C.1.2 Q 0103-294A–Q 0103-294B

These quasars belong to a group of QSOs described in Rollinde et al. (2003). There is an over-density of C IV systems between  $z_{\text{abs}} = 1.536$  and 2.18 observed in front of the group (see also Section 6.2).

#### C.1.3 Q 0236-2411–Q 0236-2413

There are strong but narrow associated absorption features toward Q 0236-2413 close to the C IV, Si IV and N V emission lines.

#### C.1.4 J 023836.9-282310–J 023849.0-282101

We detect strong metallic absorption toward QSO J 023836.9-282310 and in particular a strong O VI doublet at  $z_{\text{abs}} = 2.56$ . The signal to noise ratio of the J 023849.0-282101 spectrum is good but very few metal absorptions are

detected except for a strong Mg II system at  $z_{\text{abs}} = 0.871$ . In the same spectrum a C IV system may be present at  $z_{\text{abs}} = 2.083$  but the  $\lambda 1550$  transition is under our  $3\sigma$  detection limit.

### C.1.5 UM 680–UM 681

This pair has been observed with UVES (see D'Odorico et al. 2002). As discussed by these authors, there is a sub-DLA system ( $\log N(\text{H I}) = 18.6$ ) toward UM 681 at  $z_{\text{abs}} = 1.788$ . Noticeably as well, there are two coincident Lyman Limit Systems at  $z_{\text{abs}} = 2.03$  and two coincident associated systems at  $z_{\text{abs}} = 2.125$ .

### C.1.6 J 031036.4-305108–J 031041.0-305027

Very few metal absorptions are seen toward J 031036.4-305108 apart from a C IV system at  $z_{\text{abs}} = 1.8$ . On the contrary, the line of sight toward J 031041.0-305027 shows two metallic systems, one at  $z_{\text{abs}} \sim 2.39$  and the other clearly associated with the QSO at  $z_{\text{abs}} \sim 2.542$ . The two lines of the Si IV doublet for the latter system are under our  $3\sigma$  detection limit but weak features can be seen at the expected positions and N V absorption lines are clearly detected (see Appendix).

### C.1.7 FOCAP QSF :01–FOCAP QSF :04

There is a possible C IV doublet at  $z_{\text{abs}} = 2.27$  toward FOCAP QSF :01 but the corresponding C IV  $\lambda 1550$  transition is weaker than our detection limit.

### C.1.8 J 095800.2-002858–J 095810.9-002733

Strong C IV absorption is seen toward J 095810.9-002733 at  $z_{\text{abs}} \sim 1.807$  with other metal absorption lines from Al II, Al III and C II; the Lyman- $\alpha$  absorption corresponding to this system is out of the observed wavelength range. Possible N V  $\lambda\lambda 1238, 1242$  absorptions are seen at 4164.5 and 4151.5 Å but the corresponding features are below the  $3\sigma$  detection limit. A Si IV system may be present at  $z_{\text{abs}} = 2.122$  toward J 095800.2-002858.

### C.1.9 J 102827.1-013641–J 102832.6-013448

The determination of the emission redshift for J 102827.1-013641 is complicated by the presence of a number of absorptions at a redshift close to the emission redshift : there is a strong Lyman- $\alpha$  system with associated metallic absorption at  $z_{\text{abs}} = 2.399$ . The determination of  $z$  by a Gaussian fit of the C IV emission line gives  $z_{\text{em}} = 2.392$ . The Fe II system at  $z_{\text{abs}} = 1.316$  has no Mg II counterpart detected.

### C.1.10 Q 1206-1056–Q 1207-1057

Q 1207-1057 shows broad and shallow absorptions at 4885-4955 Å for C IV, 4385-4500 Å for S IV and 3900-3974 Å for N V. The Lyman- $\alpha$  line associated

with the BAL system is not clearly detected. There is probably a Mg II and Fe II system at  $z_{\text{abs}} \sim 0.772$  but most of the corresponding absorptions are under the  $3\sigma$  detection limit.

#### C.1.11 J 120725.9-024519—J 120734.5-024725

Strong absorptions from Mg II  $\lambda\lambda 2796, 2803$ , Fe II  $\lambda\lambda 2374, 2382$ , Fe II  $\lambda\lambda 2596, 2600$  and Mg I  $\lambda 2852$  are detected at  $z_{\text{abs}} = 0.777$  toward J 120725.9-024519.

#### C.1.12 J 123510.5-010746—J 123511.0-010830

The redshift difference between these two quasars is one of the largest in our sample :  $z_{\text{em}} = 2.785$  and  $2.235$  for J 123510.5-010746 and J 123511.0-010830 respectively. There is a  $z_{\text{abs}} \sim 2.26$  associated C IV system in the spectrum of J 123510.5-010746 at a redshift close to the emission redshift of the QSO.

#### C.1.13 J 125556.9+001848—J 125606.3+001728

The spectrum of J 125556.9+001848 presents a shallow  $z_{\text{abs}} \sim 2.08$  C IV absorption feature close to the emission redshift and a strong absorption in the range  $3810\text{--}3845 \text{ \AA}$  that could be identified as the corresponding N V absorption. J 125606.3+001728 shows several strong C IV absorptions that have no counterpart along the adjacent line of sight toward J 125556.9+001848.

#### C.1.14 J 135001.7-011703—J 135003.0-011819

A damped Lyman- $\alpha$  system is detected toward J 135003.0-011819 at  $z_{\text{abs}} = 2.33$  ( $W_{\text{obs}} = 29.6 \text{ \AA}$ ) with associated strong metallic absorption.

#### C.1.15 J 141124.6-022943—J 141117.3-023222

There is a noticeable decrease of the number of H I absorption lines in the Lyman- $\alpha$  forest of J 141124.6-022943 at the emission redshift of J 141117.3-023222 possibly corresponding to a strong transverse proximity effect.

#### C.1.16 J144245.6-024251—J144245.7-023906

Toward J 144245.6-024251, there is a damped Lyman- $\alpha$  system at  $z_{\text{abs}} = 2.218$  ( $W_{\text{obs}} = 21.4 \text{ \AA}$ ) as well as Zn II, Cr II and Fe II absorptions at  $z_{\text{abs}} = 1.178$ .

#### C.1.17 Q 2129-4653A—Q 2129-4653B

There is a strong feature in the two spectra over the wavelength range  $4044\text{--}4058 \text{ \AA}$ . In Q2129-4653B we successfully identified this feature as two blended C IV systems with a separation of about  $530 \text{ km s}^{-1}$ . Along the other line of sight the lines are heavily blended but could be modelled as C IV absorptions at the same redshift.



**C.1.18 Q 2139-4504B–Q 2139-4504A**

This is the pair with the smallest separation (0.6 arcmin) and the highest redshift in our sample. The two spectra show a Lyman limit system at a redshift close to the emission redshift of the quasar. There is a strong damped Lyman- $\alpha$  system at  $z_{\text{abs}} = 3.172$  ( $W_{\text{obs}} = 55.35 \text{ \AA}$ ) toward Q 2139-4504B with associated C II and Si II absorptions. There are no corresponding metal absorptions toward Q 2139-4504A down to  $w_r < 0.3 \text{ \AA}$ . It is interesting to note that there is a lack of absorption in the Lyman- $\alpha$  forest of Q 2139-4504B at the redshift of Q 2139-4504A suggesting the presence of a strong transverse proximity effect.

**C.1.19 J 214501.6-303121–J 214507.0-303046**

Associated systems are detected toward both quasars, at  $-1250$  and  $-380 \text{ km s}^{-1}$  relative to the QSO emission redshift toward, respectively, J 214501.6-303121 and J 214507.0-303046.

**C.1.20 J 223850.1-295612–J 223850.9-295301**

J 223850.9-295301 exhibits broad but shallow absorption lines of C IV and Lyman- $\alpha$ .

**C.1.21 J 232800.7-271655–J 232804.4-271713**

The spectra have a poor signal-to-noise ratio. We identify a possible Mg II system toward J 232800.7-271655 at  $z_{\text{abs}} = 0.368$  (Mg II  $\lambda 2803$  and Mg II  $\lambda 2796$  but no other species are seen in the spectrum. A strong C IV system may be present in the Lyman- $\alpha$  forest of J 232804.4-271713 at  $z_{\text{abs}} = 1.545$ .

## C.2 Line Lists



Q0102-2931					Q0102-293				
$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z
1	0.22	41	3943.28	CII_1334 : 1.9548	<0.45				
2	0.43	27	4114.44	SiIV_1393 1.9520	<2.49				
3	0.32	28	4117.41	SiIV_1393 1.9542	<0.91				
4	0.37	27	4140.48	SiIV_1402 1.9516	<0.27				
5	0.30	28	4144.54	SiIV_1402 1.9545	<0.64				
6	<0.43				0.74	48	4228.74	CIV_1548 1.7314	
7	<0.44				0.55	49	4235.21	CIV_1550 1.7310	
8	<0.43				0.88	44	4305.36	SiIV_1393 2.0890	
9	<0.45				1.09	40	4332.85	SiIV_1402 2.0888	
10					1.62	35	4500.45	...	...
11	<0.42				0.44	37	4547.39	CIV_1548 1.9372	
12	0.43	30	4547.44	FeII_2382 0.9085	<0.44				
13	<0.40				0.26	37	4554.88	CIV_1550 1.9372	
14	1.64	27	4563.04	CIV_1548 :1.9473	<0.33				
15	3.73	18	4570.51	CIV_1550 1.9472	<0.33				
				CIV_1548 1.952					
16	1.46	18	4573.56	CIV_1548 :1.9541	<0.32				
17	3.77	18	4576.34	CIV_1550 :1.9510	<0.32				
18	1.25	24	4581.03	CIV_1550 :1.9540	<0.32				
19	<0.42				1.39	38	4782.31	CIV_1548 2.0889	
20	<0.43				0.86	40	4790.25	CIV_1550 2.0889	
23	0.48	29	4866.89	CIV_1548 2.1436	<0.30				
24	0.30	30	4874.66	CIV_1550 2.1434	<0.30				
25	0.35	33	4902.52	CIV_1548 2.1666	<0.31				
26	0.24	35	4910.43	CIV_1550 2.1664	<0.31				
27	0.20	39	4935.25	FeII_2586 0.9080	<0.32				
28	0.11	46	4960.91	FeII_2600 0.9079	<0.32				
29	<0.28				0.48	38	5015.98	CIV_1548 2.2399	
30	<0.30				0.18	38	5024.29	CIV_1550 2.2398	
31	<0.34				2.11	34	5071.72	FeII_2344 1.1635	
32	<0.37				1.48	36	5137.20	FeII_2374 1.1635	
33	<0.38				2.46	33	5155.22	FeII_2382 1.1635	
34	0.58	30	5335.73	MgII_2796 0.9081	<0.25				
35	0.52	31	5349.39	MgII_2803 0.9081	<0.25				
38	<0.46				2.06	33	5596.51	FeII_2586 1.1636	
39	<0.47				2.46	33	5625.82	FeII_2600 1.1636	

Q0103-294A					Q0103-294B				
$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z
1	0.13	63	3904.00	...	...				
3	<0.20				0.37	49	3925.59	CIV_1548 1.5356	
4	<0.20				0.26	48	3932.93	CIV_1550 1.5361	
5	<0.20				0.44	47	3941.42	CIV_1548 1.5458	
6	<0.20				0.22	49	3947.88	CIV_1550 1.5457	
7	<0.20				1.76	43	3956.64	CIV_1548 :1.5556	
8	<0.20				1.25	46	3963.38	CIV_1550 :1.5557	
9	2.84	50	4001.65	...					
10	<0.25				0.15	42	4061.08	SiIV_1393 1.9138	
11	<0.26				0.13	41	4087.95	SiIV_1402 1.9142	
12					0.53	42	4187.52	...	...
13	1.28	44	4238.75	CII_1334 2.1762	<0.29				
14	1.93	42	4270.88	CIV_1548 1.7586	<0.28				
15	1.23	44	4277.66	CIV_1550 1.7584	<0.28				
16	<0.25				0.98	42	4373.40	SiIV_1393 2.1378	
17	<0.24				0.56	43	4401.87	SiIV_1402 2.1380	
18	0.83	49	4426.89	SiIV_1393 2.1762	<0.27				
19	0.46	53	4455.40	SiIV_1402 2.1761	<0.26				
20	0.16	54	4484.33	...					
21	0.23	55	4484.34	CIV_1548 :1.8965	<0.26				
22	0.08	55	4500.99	CIV_1550 :1.9024	<0.25				
23	<0.23				0.90	45	4511.06	CIV_1548 1.9137	
24	<0.23				0.51	45	4518.63	CIV_1550 1.9138	
25	0.28	50	4566.26	CIV_1548 1.9494	<0.27				
26	0.27	50	4574.08	CIV_1550 1.9495	<0.28				
29	<0.25				0.69	44	4724.32	CIV_1548 2.0515	
30	<0.26				0.37	45	4732.09	CIV_1550 2.0514	
31	0.17	46	4759.04	...					
32	0.45	48	4769.66	...					
35	<0.22				2.70	41	4857.95	CIV_1548 2.1378	
36	<0.21				2.14	44	4865.95	CIV_1550 2.1377	
37	1.40	61	4918.35	CIV_1548 2.1768	<0.22				
38	0.42	61	4921.98	CIV_1548 :2.1792	<0.22				
39	0.68	61	4924.78	CIV_1550 2.1757	<0.21				
40	0.37	65	4929.78	CIV_1550 :2.1789	<0.21				

J 013734.2-303802					J 013734.2-304050				
$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z
1	<1.00				0.33	58	4128.70	CII_1334 2.0937	
2					0.33	44	4216.09	...	...
5	0.46	40	4291.65	...					
6	0.47	41	4321.63	NV_1238 2.4885	<0.31				
7	0.21	40	4335.65	NV_1242 2.4886	<0.31				
9	0.32	37	4353.71	...					
10	<0.40				0.27	39	4500.24	SiIV_1393 2.2289	
11	<0.40				0.35	40	4529.05	SiIV_1402 2.2286	
12	<0.41				0.40	44	4636.20	CIV_1548 1.9946	
13	<0.40				0.30	45	4643.51	CIV_1550 1.9943	
14	<0.40				0.62	39	4790.69	CIV_1548 2.0944	
15	<0.39				0.41	40	4798.57	CIV_1550 2.0943	
18	<0.36				0.65	39	4849.10	CIV_1548 2.1321	
19	<0.36				0.41	39	4856.68	CIV_1550 2.1318	
20	0.96	33	4861.62	SiIV_1393 2.4881	<0.31				
21	0.22	34	4873.05	...					
22	0.54	33	4892.82	SiIV_1402 2.4880	<0.31				
23	<0.40				1.16	40	4999.07	CIV_1548 2.2289	
24	<0.40				0.89	41	5007.56	CIV_1550 2.2291	
25	0.32	32	5239.63	...					
26	2.34	50	5400.30	CIV_1548 2.4881	<0.29				
27	1.86	48	5409.27	CIV_1550 2.4881	<0.29				

Q0236-2411					Q0236-2413				
$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z
1	<1.37				5.87	25	3947.47	NV_1238 2.1865	
2	<0.60				6.31	24	3960.15	NV_1242 2.1865	
3	<1.30				2.69	34	3971.02	NV_1238 2.2055	
4	<1.44				1.72	38	3984.24	NV_1242 2.2058	
6					0.53	35	4007.95	...	...
7	0.82	41	4046.73	NV_1238 2.2666	<0.38				
8	0.57	40	4059.44	NV_1242 2.2664	<0.38				
9					1.26	31	4198.12	...	...
10	<0.37				0.34	30	4277.66	CII_1334 2.2054	
11	0.67	30	4326.23	CIV_1548 1.7944	<0.41				
12	<0.58				0.32	30	4329.38	CIV_1548 1.7964	
13	0.52	30	4333.22	CIV_1550 1.7942	<0.41				
14	<0.40				0.24	30	4336.91	CIV_1550 1.7966	
15	<0.37				1.88	28	4374.74	CIV_1548 1.8257	
16	<0.36				1.20	28	4382.07	CIV_1550 1.8257	
17	<0.38				0.97	28	4440.65	SiIV_1393 2.1861	
18	<0.37				1.81	31	4468.34	SiIV_1393 2.2060	
19	<0.37				1.81	31	4468.34	SiIV_1402 2.1854	
20	<0.37				1.41	34	4497.10	SiIV_1402 2.2059	
21					0.86	33	4798.68	...	...
23	<1.11				0.92	32	4842.83	CIV_1548 2.1280	
26	<0.35				0.61	33	4850.33	CIV_1550 2.1277	
27					0.49	31	4903.63	...	...
28					0.75	32	4914.88	...	...
29	<0.33				5.81	24	4933.50	CIV_1548 2.1866	
30	<0.33				5.28	24	4939.86	CIV_1548 2.1907	
31	<0.33				5.28	24	4939.86	CIV_1550 2.1854	
32	<0.33				1.37	36	4949.44	CIV_1550 2.1916	
33	<0.33				3.65	36	4962.58	CIV_1548 2.2054	
34	<0.32				3.22	48	4971.20	CIV_1550 2.2056	
35	0.64	39	4989.17	CIV_1548 2.2226	<0.23				
36	0.18	41	4997.53	CIV_1550 2.2226	<0.27				
37	0.41	47	5057.29	CIV_1548 2.2666	<0.33				
38	0.40	47	5065.72	CIV_1550 2.2666	<0.33				
39	0.35	37	5144.38	...					
40	0.61	36	5229.01	...					
43	0.76	34	5706.23	...					

J 023836.9-282310					J 023849.0-282101						
	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	$z$		$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	$z$
1	4.02	22	3671.50	OVI_1031	2.5579		<0.28				
2	3.50	28	3691.45	OVI_1037	2.5576		<0.24				
4							0.28	84	4385.94	...	...
5	0.53	59	4386.18	FeII_2382	0.8408		<0.25				
6	4.61	50	4406.52	SiIV_1393	2.1616		<0.15				
7	4.61	50	4406.52	NV_1238	2.5570		<0.15				
				SiIV_1393	2.162						
8	2.49	52	4421.26	NV_1242	2.5575		<0.15				
9	0.79	56	4434.20	SiIV_1402	2.1610		<0.15				
10	<0.25						0.42	84	4457.81	FeII_2382	0.8709
11	<0.27						0.40	85	4477.99	CIV_1548	1.8924
12	<0.27						0.26	86	4485.41	CIV_1550	1.8924
15							0.22	97	4538.02	...	...
16	0.92	38	4547.64	CIV_1548	1.9374		<0.13				
18	0.47	39	4555.63	CIV_1550	1.9376		<0.13				
19							0.17	87	4625.12	...	...
20	0.63	39	4740.89	...	...		<0.15				
21	0.43	41	4760.69	FeII_2586	0.8405		0.39	85	4773.10	...	...
22							<0.14				
23	0.37	41	4784.57	FeII_2600	0.8401		0.14	87	4839.11	FeII_2586	0.8708
25	<0.31						0.50	87	4864.79	FeII_2600	0.8710
26	<0.30						<0.38				
27	4.48	32	4894.12	CIV_1548	2.1612		0.31	89	4896.23	CIV_1548	2.1625
28	<3.02						<0.17				
29	1.67	33	4901.45	CIV_1550	2.1606		0.22	90	4904.00	CIV_1550	2.1623
30	<1.42										
31	0.26	41	5042.31	...	...		<0.13				
32	0.92	38	5146.51	MgII_2796	0.8404		<0.14				
33	0.62	40	5159.36	MgII_2803	0.8403						
34	0.68	39	5173.96	...	...		2.42	84	5231.69	MgII_2796	0.8709
37	<0.31						1.63	85	5245.00	MgII_2803	0.8709
38	<0.30										
41	0.17	40	5329.20	...	...		<0.15				
46	2.53	50	5507.65	CIV_1548	2.5574		<0.15				
47	1.88	53	5515.96	CIV_1550	2.5569						
53	0.62	38	5741.97	...	...						

J 031036.4-305108						J 031041.0-305027						J 095800.2-002858						J 095810.9-002733					
$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z		$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z		$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z		
1	1.11	17	3736.42	CII_1334	1.7998	<0.62						1	<1.69					5.47	10	3745.16	CII_1334	1.8063	
2	0.68	52	4334.83	CIV_1548	1.7999	<0.20															HI_1215	2.0807	
3	0.78	48	4341.97	CIV_1550	1.7999	<0.21												2.83	16	3937.55	SiIV_1402	1.8070	
4	0.28	40	4380.07	...	...																HI_1215	2.2390	
5	<0.31					0.81	51	4388.24	NV_1238	2.5423		3	2.22	31	4144.25	CIV_1548	1.6768	<3.98					
6	0.48	38	4393.47	NV_1238	2.5465	<0.24						4	2.24	30	4151.67	CIV_1550	1.6771	<1.03					
7	<0.32					0.47	49	4402.48	NV_1242	2.5424		5	0.80	33	4166.82	CII_1334	2.1223	<0.63					
8	0.54	38	4407.88	NV_1242	2.5467	<0.25						6	1.42	18	4338.31	SiIV_1393	2.1127	<0.30					
9						0.28	48	4412.16	...	...		7	<1.35					6.88	35	4345.74	CIV_1548	1.8070	
12	<0.43					0.65	33	4649.20	MgII_2796	0.6626		8	1.44	19	4351.30	SiIV_1393	2.1220	<4.05					
13	<0.45					0.62	33	4661.19	MgII_2803	0.6626		9	<0.91					4.60	31	4352.28	CIV_1550	1.8065	
14	0.37	26	4677.29	AIII_1670	1.7995	<0.37						10	0.84	20	4366.54	SiIV_1402	2.1128	<0.23					
15	<0.47					1.12	32	4725.06	SiIV_1393	2.3902		11	0.58	21	4379.53	SiIV_1402	2.1221	<0.24					
16	<0.46					0.57	33	4754.80	SiIV_1402	2.3896		12	<0.55					2.40	32	4460.73	SiIV_1393	2.2005	
17	0.77	26	5079.98	...	...	<0.39						13	<0.52					1.09	33	4490.93	SiIV_1402	2.2015	
18	0.85	26	5193.08	AIII_1854	1.7999	<0.38						14	<0.52					1.56	33	4491.24	SiIV_1393	2.2224	
19	0.69	27	5215.09	AIII_1862	1.7996	2.89	27	5247.96	CIV_1548	2.3897		15	<0.52					0.79	32	4528.58	SiIV_1402	2.2283	
20	<0.45					2.35	26	5256.67	CIV_1550	2.3897		16	<0.44					4.30	27	4687.29	AIII_1670	1.8054	
21	<0.45					0.94	46	5484.23	CIV_1548	2.5423		17	0.24	25	4765.54	...	...	<0.38					
22	<0.31					0.95	46	5494.10	CIV_1550	2.5428		18	2.91	19	4819.77	CIV_1548	2.1131	<0.39					
23	<0.70					<0.51						19	1.68	17	4827.44	CIV_1550	2.1129	<0.37					
24	1.13	42	5499.14	CIV_1548	2.5519	<0.26						20	2.83	18	4832.54	CIV_1548	2.1214	<0.37					
25	0.32	45	5507.67	CIV_1550	2.5515							21	1.57	20	4842.14	CIV_1550	2.1224	<0.37					
												22	1.10	22	4855.92	...	...						
												24	<0.51					0.74	31	4869.18	CIV_1548	2.1450	
												25	<0.50					0.32	32	4878.13	CIV_1550	2.1456	
												27	<0.51					1.39	32	4954.16	CIV_1548	2.1999	
												28	<0.50					0.43	35	4963.93	CIV_1550	2.2009	
												29	<0.49					2.62	30	4988.67	CIV_1548	2.2222	
												30	<0.50					1.54	32	4996.40	CIV_1550	2.2219	
												33	0.44	40	5198.17	...	...						
												34	<0.30					1.52	30	5206.14	AIII_1854	1.8070	
												35	0.98	40	5216.86	CIV_1548	2.3696	<0.37					
												36	0.37	43	5224.94	CIV_1550	2.3692	<0.44					
												37	<0.38					0.92	32	5229.07	AIII_1862	1.8071	
												38						1.09	37	5384.46	...	...	

J 111200.4-013242					J 111201.8-013018					Q 1206-1056					Q 1207-1057					
$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	
1	<0.81				1.64	43	3933.49	SiIV_1393	1.8222	1	0.28	58	4064.12	...						
2	<1.72				1.06	50	3959.22	SiIV_1402	1.8224	2	0.42	44	4152.90	SiIV_1393	1.9796	<0.74				
3	0.24	50	4127.70	...						3	0.24	41	4178.74	SiIV_1402	1.9789	<0.76				
4	0.32	45	4151.18	CIV_1548	1.6813	<0.74				4	<0.34					5.30	50	4236.43	NV_1238	2.4197
5	0.48	46	4157.79	CIV_1550	1.6811	<0.21				5	0.97	34	4244.90	SiIV_1393	2.0456	<3.32				
6	0.27	39	4194.57	CIV_1548	1.7093	<0.40				6	<0.35					3.95	49	4249.23	NV_1242	2.4191
7	0.25	40	4202.22	CIV_1550	1.7097	<0.52				7	0.46	35	4272.25	SiIV_1402	2.0456	<0.20				
9	<0.31				1.35	82	4362.97	CIV_1548	1.8181	8						0.28	60	4276.06	...	...
10	<0.31				0.78	81	4370.12	CIV_1550	1.8180	9	<0.34					1.02	57	4289.60	NV_1238	2.4626
11	<0.30				0.68	80	4396.84	NV_1238	2.5492	10	<0.34					1.25	54	4302.66	NV_1242	2.4621
12	<0.29				0.23	82	4404.75	NV_1242	2.5442	11						0.30	51	4335.88	...	...
13	<0.29				0.90	63	4494.26	CIV_1548	1.9029	12	0.97	40	4611.88	CIV_1548	1.9789	<0.27				
14	<0.29				0.56	62	4501.22	CIV_1550	1.9025	13	0.33	42	4619.99	CIV_1550	1.9791	<0.26				
15	0.37	43	4535.94	...						14						0.34	45	4712.18	...	...
16	0.96	50	4616.26	...						15	1.90	34	4715.08	CIV_1548	2.0455	<0.32				
17	0.54	46	4646.14	...						16	1.25	34	4723.00	CIV_1550	2.0456	<0.26				
18					0.21	59	4671.42	...	...	17						1.28	46	4750.74	...	...
19	0.64	41	4676.47	...						18	0.91	35	4779.70	CIV_1548	2.0873	<0.25				
21					0.32	57	4716.14	...	...	19	0.51	35	4787.98	CIV_1550	2.0875	<0.25				
22	0.67	38	4866.52	CIV_1548	2.1433	<0.21				20						0.36	50	4793.47	...	...
23	0.40	39	4873.27	CIV_1550	2.1425	<0.21				21	<0.34					1.19	51	4835.93	SiIV_1393	2.4697
24	0.82	37	4881.60	MgII_2796	0.7457	<0.20				22						0.49	51	4857.82	...	...
25					0.14	60	4890.22	...	...	23	<0.34					0.54	51	4866.50	SiIV_1402	2.4692
26	0.43	39	4894.66	MgII_2803	0.7459	<0.20				24	1.50	36	4954.54	MgII_2796	0.7718	<0.26				
27	<0.29				0.19	67	4972.09	SiIV_1393	2.5674	25	0.92	37	4967.29	MgII_2803	0.7718	<0.26				
28	0.57	45	5038.37	CIV_1548	2.2543	<0.20				26						0.55	46	5163.93	...	...
29	0.42	47	5047.34	CIV_1550	2.2547	<0.20				30						1.37	49	5227.09	...	...
30	0.24	56	5105.27	...						31	<0.30					1.34	49	5247.77	CIV_1548	2.3896
31	<0.29				1.57	53	5205.01	CIV_1548	2.3620	32	<0.30					0.77	50	5258.31	CIV_1550	2.3907
32	<0.30				1.14	53	5213.72	CIV_1550	2.3620	33						0.79	50	5275.34	...	...
33	<0.31				0.28	55	5283.52	CIV_1548	2.4127	34	<0.31					2.43	44	5297.09	CIV_1548	2.4214
34	<0.30				0.12	55	5292.33	CIV_1550	2.4127	35	<0.30					1.64	47	5302.08	CIV_1550	2.4190
36	<0.30				0.61	66	5488.32	CIV_1548	2.5450	36						0.26	50	5313.11	...	...
37	<0.30				0.57	69	5495.51	CIV_1550	2.5437	37	<0.31					0.97	50	5360.70	CIV_1548	2.4625
38	<0.30				0.79	68	5523.19	CIV_1548	2.5675	38	<0.31					5.16	40	5371.25	CIV_1550	2.4636
39	<0.31				0.44	66	5532.46	CIV_1550	2.5675									CIV_1548	2.4694	
43					0.58	53	5731.16	...	...	39	<0.31					2.25	45	5381.44	CIV_1550	2.4701

J 112108.2+003420					J 112116.1+003112					J 123510.5-010746					J 123511.0-010830				
$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z
1	0.32	46	3935.02	...						1	<0.48				1.44	15	3627.43	SiIV_1393	1.6026
2					0.51	31	4019.10	...	...	2	<0.42				1.42	18	3650.60	SiIV_1402	1.6024
3	0.40	36	4065.07	...						3	<0.24				1.08	50	3946.92	CIV_1548	1.5494
4	0.58	34	4085.84	CII_1334	2.0616	<0.42				4	<1.44				0.78	45	3953.77	CIV_1550	1.5495
5					0.65	27	4125.55	...	...	5					1.00	42	3974.10	...	...
6					0.58	25	4246.95	...	...	6	<0.24				2.57	35	4029.43	CIV_1548	1.6027
7	1.23	34	4267.73	SiIV_1393	2.0620	<1.32				7	<0.30				2.34	34	4036.12	CIV_1550	1.6026
8					1.89	26	4279.14	...	...	8					2.99	32	4193.30	...	...
9	1.14	33	4295.44	SiIV_1402	2.0621	<0.44				9					0.37	37	4318.36	...	...
10					0.31	30	4464.20	...	...	10	<1.28				1.34	32	4348.96	AlII_1670	1.6029
11	0.53	38	4492.89	...						11	<0.19				2.23	32	4797.41	CIV_1548	2.0987
12	<0.34				1.00	27	4553.79	CIV_1548	1.9413	12	<0.20				1.64	34	4805.60	CIV_1550	2.0988
13	<0.34				0.68	27	4561.16	CIV_1550	1.9412	13	<0.21				0.94	36	4827.95	AlIII_1854	1.6031
14	0.66	35	4660.61	...						14	<0.21				0.63	38	4848.76	AlIII_1862	1.6030
15	<0.35				0.85	26	4668.46	CIV_1548	2.0154	15	0.24	57	4889.18	...					
16	<0.35				0.70	27	4675.78	CIV_1550	2.0151	16	1.12	53	5015.39	CIV_1548	2.2395	<0.23			
17	<0.34				1.46	26	4717.58	CIV_1548	2.0471	17	0.60	54	5023.48	CIV_1550	2.2393	<0.24			
18	<0.34				1.16	27	4725.59	CIV_1550	2.0472	18	0.29	58	5080.15	CIV_1548	2.2813	<0.31			
19	1.54	34	4741.18	CIV_1548	2.0624	<0.43				19	0.30	57	5088.41	CIV_1550	2.2812	<0.32			
20	1.51	33	4748.41	CIV_1550	2.0619	<0.43				20	0.29	56	5119.20	...					
21					0.44	28	4793.89	...	...	21	0.23	57	5132.99	...					
22	<0.27				1.97	46	4976.65	CIV_1548	2.2145										
23	<0.28				0.97	43	4985.13	CIV_1550	2.2146										
25					0.39	30	5374.31	...	...										
26	<0.34				0.86	28	5407.04	MgII_2796	0.9336										
27	<0.34				0.50	28	5419.57	MgII_2803	0.9331					</					

J 135001.7-011703						J 135003.0-011819						J 144245.6-024251						J 144245.7-023906					
$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z		$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z		$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z		$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	
1	2.28	51	3863.44	MgII_2796	0.3816	<0.87						1	1.30	21	3830.49	SiII_1190	2.2178	<0.51					
2	1.04	56	3872.57	MgII_2803	0.3813	<0.93						2	1.45	20	3840.14	SiII_1193	2.2181	<0.52					
3	0.40	43	3908.55	...	...							3	4.24	18	3881.79	SiIII_1206	2.2174	<0.51					
4	0.47	43	3937.48	NV_1238	2.1784	<0.80																	
5	0.37	45	3950.47	NV_1242	2.1787	<2.36						4	2.99	47	4055.69	SiII_1260	2.2177	<1.73					
6	<0.31					4.13	11	3969.05	SiII_1190	2.3342													
7	<0.34					5.85	10	3980.30	SiII_1193	2.3356													
8	3.05	26	4019.18	CIV_1548	1.5960	<2.57						5	2.04	35	4190.22	OI_1302	2.2179	<3.60					
9	<3.20					5.60	9	4023.80	SiIII_1206	2.3351		6	0.93	35	4197.33	SiII_1304	2.2179	<1.29					
10	1.77	27	4024.81	CIV_1550	1.5953	<2.91						7	2.42	26	4294.85	CII_1334	2.2182	<0.46					
11	<0.43					4.66	9	4203.55	SiII_1260	2.3350		8	0.57	33	4354.19	CIV_1548	1.8124	<0.24					
14	0.67	28	4329.50	...	...							9	0.27	34	4361.32	CIV_1550	1.8123	<0.26					
15	<0.42					4.15	10	4342.08	OI_1302	2.3345		10	0.73	31	4412.47	ZnII_2026	1.1778	<0.29					
16	<0.42					3.42	18	4350.76	SiII_1304	2.3355		11	0.26	34	4477.62	CrII_2056	1.1776	<0.36					
17	0.48	30	4357.76	...	...							12	2.08	31	4486.38	SiIV_1393	2.2189	<0.36					
18	<0.50					8.81	42	4451.15	CII_1334	2.3354		13	0.51	32	4491.41	ZnII_2062	1.1775	<0.36					
20	0.53	37	4465.46	...	...							14	0.78	30	4514.82	SiIV_1402	2.2185	<0.37					
24						0.48	21	4623.58	...	...		15	<0.31					0.86	32	4663.84	CIV_1548	2.0124	
25	<0.40					4.19	17	4648.85	SiIV_1393	2.3355		16	<0.31					0.23	31	4672.11	CIV_1550	2.0127	
26	<0.40					1.75	18	4679.18	SiIV_1402	2.3357		17	0.20	36	4717.52	...	...						
27	0.83	31	4792.42	CIV_1548	2.0955	<0.54						18	0.35	35	4724.86	...	...	<0.38					
28	0.27	32	4799.76	CIV_1550	2.0951	<0.54						19	1.44	31	4912.86	SiII_1526	2.2179	1.00	34	4971.21	SiIV_1393	2.5668	
29	1.92	30	4809.41	CIV_1548	2.1064	<0.58						20	<0.35					<0.34					
30	1.56	30	4816.95	CIV_1550	2.1061	<0.59						21	2.23	32	4983.29	CIV_1548	2.2188	<0.35					
31						0.86	20	4858.18	...	...		22	1.27	33	4991.51	CIV_1550	2.2187	0.51	34	5003.26	SiIV_1402	2.5667	
32	0.56	40	4864.81	...	...							23	<0.34					<0.38					
34						2.22	20	4910.26	...	...		24	2.84	36	5104.77	FeII_2344	1.1776	<0.38					
35	0.37	53	4921.56	CIV_1548	2.1789	<0.60						28	2.46	45	5171.35	FeII_2374	1.1779	<0.38					
36	0.17	56	4929.84	CIV_1550	2.1789	<0.57						29	3.69	43	5188.78	FeII_2382	1.1776	<0.38					
39	<0.35					3.79	18	5091.70	SiII_1526	2.3351		35	<0.33					0.73	31	5282.85	CIV_1548	2.4122	
40	<0.35					4.74	11	5166.22	CIV_1548	2.3369		36	<0.33					0.48	32	5289.71	CIV_1550	2.4110	
41	<0.36					2.93	16	5173.72	CIV_1550	2.3362		40						0.82	32	5309.69	...	...	
42	<0.38					2.71	17	5364.37	FeII_1608	2.3351		46						0.95	32	5339.24	...	...	
43	<0.39					0.98	19	5374.09	FeII_1611	2.3355		52	1.65	34	5376.62	AlII_1670	2.2180	<0.37					
44						3.13	19	5454.58	...	...		58	<0.33					0.46	49	5513.18	CIV_1548	2.5610	
45	<0.38					1.34	17	5569.91	AlII_1670	2.3337		59	<0.36					2.14	43	5521.69	CIV_1550	2.5606	
46	<0.40					0.65	30	5639.95	CIV_1548	2.6429		62	<0.33								CIV_1548	2.5665	
47	<0.40					0.37	32	5648.92	CIV_1550	2.6426		65	0.54	35	5611.89	...	...	1.66	44	5531.15	CIV_1550	2.5667	
48	<0.40					1.68	39	5671.44	CIV_1548	2.6632		66	3.40	31	5632.77	FeII_2586	1.1776	<0.37					
49	<0.40					1.36	39	5681.35	CIV_1550	2.6635		67	3.90	28	5662.22	FeII_2600	1.1776	<0.37					
												68	0.33	34	5676.30	...	...						
												70	0.38	36	5710.42	...	...						
J 141124.6-022943						J 141117.3-023222						Q2129-4653A						Q2129-4653B					
$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z		$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z		$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z		$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	
1	<2.11					0.87	65	4104.61	NV_1238	2.3133		1	0.42	39	3916.03	SiIV_1393	1.8097	<1.24					
2	<6.53					1.18	63	4117.01	NV_1242	2.3127		2	0.25	34	3933.37	...	...						
3	<6.53					1.17	64	4117.01	SiIV_1393	1.9539		3	0.33	34	3941.81	SiIV_1402	1.8100	<0.23					
4	<0.29					0.52	58	4142.86	SiIV_1402	1.9533		4	0.31	34	3990.91	SiII_1526	1.6141	<0.27					
5	<1.33					1.25	48	4386.90	SiIV_1393	2.1475		5	<0.94					0.51	36	4044.41	CIV_1548	1.6123	
6	<0.86					0.87	50	4415.04	SiIV_1402	2.1474		6	1.36	24	4047.24	CIV_1548	1.6142	<0.45					
7	<0.16					1.62	50	4572.28	CIV_1548	1.9533		7	1.01	24	4051.70	CIV_1550	1.6127	<0.45					
8	<0.16					0.98	52	4580.08	CIV_1550	1.9534		8	<1.52					1.21	33	4051.84	CIV_1550	1.6128	
11	0.27	58	4635.05	CIV_1548	1.9938	<0.21														CIV_1548	1.6171		
12	<0.29					0.37	57	4640.29	CIV_1548	1.9972		9	0.65	23	4052.14	CIV_1548	1.6173	<0.42					
13	0.15	58	4644.79	CIV_1550	1.9951	<0.22						10	0.28	26	4055.71	CIV_1550	1.6153	<0.38					
14	<0.22					0.23	58	4647.60	CIV_1550	1.9969		11	<0.50					0.56	34	4058.62	CIV_1550	1.6171	
15	1.79	46	4718.27	MgII_2796	0.6873	<0.23						12	1.14	23	4186.23	CIV_1548	1.7039	<0.40					
16	1.40	48	4730.79	MgII_2803	0.6874	<0.24						13	0.62	24	4193.46	CIV_1550	1.7041	<0.39					
17	0.41	49	4769.57	CII_1334	2.5740	<0.23						14	0.51	24	4290.52	SiII_1526	1.8103	<0.39					
18	<0.24					2.42	44	4873.32	CIV_1548	2.1477		15	<0.49					0.83	31	4327.13	CIV_1548	1.7949	
19	<0.24					2.15	47	4881.34	CIV_1550	2.1477		16											

Q2139-4504A					Q2139-4504B				
$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z
1	<2.25				8.02	12	4320.48	CII_1036	3.1690
2	<0.34				8.29	12	4962.75	SiII_1190	3.1689
								HI_1215	3.0823
3	<0.40				5.51	16	4973.07	SiII_1193	3.1675
								HI_1215	3.0908
6	1.59	31	5001.80	SiIV_1393	<4.93				
7	0.80	34	5024.51	SiIV_1402	<6.61				
9	<0.35				3.19	16	5027.95	SiIII_1206	3.1674
10	0.82	20	5130.70	...					
11					0.34	46	5224.46	...	...
12					2.33	40	5234.87	...	...
13	0.83	19	5240.81	...					
14	<0.69				8.75	26	5253.21	SiII_1260	3.1678
15	0.81	21	5286.11	...					
16	<0.65				0.87	34	5355.71	SiII_1260	3.2491
20	<0.67				6.52	18	5427.63	OI_1302	3.1681
21	<0.68				3.44	24	5436.45	SiII_1304	3.1679
25	1.19	17	5557.65	CIV_1548	<4.26				
26	<2.11				8.94	19	5562.47	CII_1334	3.1681
								HI_1215	3.5756
27	1.01	18	5567.12	CIV_1550	<3.55				
28	1.20	20	5604.82	...					
29	<0.55				1.04	28	5670.52	CII_1334	3.2491
30	0.74	21	5688.60	...					
31	0.95	20	5704.82	...					
J 214501.6-303121					J 214507.0-303046				
$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z
1	<0.61				3.99	16	3641.56	OVI_1031	2.5289
2	<0.57				4.03	23	3661.77	OVI_1037	2.5290
3	0.36	57	3956.66	...					
4	0.60	56	3968.16	...					
5	0.23	55	3980.43	...					
6	0.33	55	3991.35	...					
7	0.44	51	4002.21	...					
8	<0.28				1.37	39	4087.01	SiIV_1393	1.9324
								HI_1215	2.3619
9	<0.29				0.43	42	4113.24	SiIV_1402	1.9322
								HI_1215	2.3835
10	0.25	42	4204.39	CIV_1548	<0.26				
11	0.26	43	4210.65	CIV_1550	<0.45				
12	0.45	42	4305.48	...					
13	<0.28				3.62	57	4371.73	NV_1238	2.5289
14	<0.28				2.65	60	4385.68	NV_1242	2.5289
16	0.21	47	4464.60	SiIV_1393	<0.23				
17	0.11	49	4494.86	SiIV_1402	<0.24				
18	<0.26				3.90	40	4540.16	CIV_1548	1.9325
19	<0.26				2.37	41	4547.69	CIV_1550	1.9325
20	<0.24				0.98	51	4888.98	CIV_1548	2.1578
21	<0.24				0.57	53	4897.20	CIV_1550	2.1579
22	0.25	52	4901.73	...					
23	<0.22				0.50	56	4917.76	SiIV_1393	2.5284
24	<0.20				0.25	58	4949.59	SiIV_1402	2.5284
25	1.46	60	4961.46	CIV_1548	<0.21				
26	0.75	63	4969.77	CIV_1550	<0.21				
32	<0.28				2.72	56	5463.43	CIV_1548	2.5289
34	<0.28				2.92	60	5471.99	CIV_1550	2.5285
J 215225.8-283058					J 215240.0-283251				
$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z
1	<3.33				2.20	22	3817.99	OVI_1031	2.6999
2	<0.72				1.33	27	3839.42	OVI_1037	2.7002
3	<2.08				2.55	38	4523.59	SiIV_1393	2.2456
4	<0.15				1.00	41	4552.89	SiIV_1402	2.2456
5	0.30	59	4579.00	...					
6	<0.33				0.83	41	4582.40	CIV_1548	1.9598
7	0.34	58	4588.50	...					
8	<0.34				0.47	42	4589.52	CIV_1550	1.9595
11	1.30	47	4681.66	FeII_2344	<0.37				
12					1.31	32	4686.15	...	...
13	0.80	47	4691.48	CIV_1548	<0.39				
14	0.54	47	4698.61	CIV_1550	<0.38				
15	0.76	45	4741.79	FeII_2374	<0.40				
16	1.66	42	4758.26	FeII_2382	<0.41				
17	0.56	43	4918.49	CIV_1548	<0.42				
18	0.28	43	4926.51	CIV_1550	<0.42				
19	<0.28				3.45	23	5025.22	CIV_1548	2.2458
20	<0.29				2.44	25	5033.55	CIV_1550	2.2458
21	0.42	42	5070.70	...					
22	0.50	43	5153.70	...					
25	0.99	44	5165.54	FeII_2586	<0.37				
26	1.84	44	5192.53	FeII_2600	<0.42				
29	<0.29				0.71	28	5423.13	AlII_1670	2.2459
30	<0.29				1.33	26	5468.44	CIV_1548	2.5321
31	<0.29				0.73	27	5477.24	CIV_1550	2.5319
32	0.16	41	5539.90	...					
33	2.94	33	5583.49	MgII_2796	<0.43				
34	3.26	35	5598.24	MgII_2803	<0.41				
35	0.75	41	5623.79	...					
36	<0.28				1.23	30	5682.18	CIV_1548	2.6702
37	<0.27				0.89	31	5691.53	CIV_1550	2.6701
38	0.51	43	5697.65	...					
39	<0.25				1.61	32	5728.49	CIV_1548	2.7001
40	<0.24				0.32	33	5739.57	CIV_1550	2.7011
J 223850.1-295612					J 223850.9-295301				
$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z
1	0.62	62	4270.70	...					
2	1.47	57	4293.80	...					
3	0.38	49	4343.59	...					
5	0.53	46	4348.16	...					
6	0.79	45	4364.05	CIV_1548	<0.26				
7	0.62	44	4370.17	CIV_1550	<0.26				
8	1.38	39	4404.23	CIV_1548	<0.27				
9	0.73	40	4411.37	CIV_1550	<0.27				
10	0.97	40	4469.96	...					
11	1.41	40	4485.15	SiIV_1393	<0.25				
12	0.95	41	4514.36	SiIV_1402	<0.25				
13	<0.33				0.46	50	4518.18	CIV_1548	1.9183
14	<0.27				0.18	51	4526.01	CIV_1550	1.9185
18	1.09	42	4621.54	...					
19	0.29	42	4666.25	...					
22	<0.27				1.87	46	4793.34	CIV_1548	2.0961
23	<0.27				1.51	48	4800.82	CIV_1550	2.0957
24	0.71	46	4808.78	...					
25	0.31	47	4820.37	...					
26	0.25	47	4840.12	...					
27	0.46	47	4850.49	CIV_1548	<0.27				
28	0.35	47	4858.65	CIV_1550	<0.28				
29	0.66	46	4869.56	CIV_1548	<0.30				
30	0.27	45	4877.67	CIV_1550	<0.32				
31					0.29	37	4889.18	...	...
32	<0.28				1.14	39	4920.49	CIV_1548	2.1782
33	<0.28				0.66	42	4929.19	CIV_1550	2.1785
34	2.62	37	4981.71	CIV_1548	<0.59				
35					0.97	50	4983.52	...	...
36	1.67	38	4989.91	CIV_1550	<0.28				
37					1.49	49	5016.96	...	...
38					0.22	56	5147.22	...	...
J 230301.6-290027					J 230318.4-290120				
$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z	$w_{\text{obs}}$	S/N	$\lambda_{\text{obs}}$	Ident	z
1	0.97	47	4430.38	CIV_1548	<0.21				
2	0.55	46	4438.40	CIV_1550	<0.21				
3	<0.27				0.43	57	4444.72	NV_1238	2.5879
4	<0.29				0.33	57	4459.40	NV_1242	2.5882
5	0.52	43	4464.84	CIV_1548	<0.22				
6	0.26	42	4472.66	CIV_1550	<0.23				
7	0.18	33	4802.82	...					
9	0.49	34	4908.45	...					
10	<0.32				0.68	42	4973.79	FeII_2344	1.1217
11	<0.34				1.07	41	5055.85	FeII_2382	1.1218
12	0.41	34	5069.16	CIV_1548	<0.29				
13	0.52	34	5077.52	CIV_1550	<0.29				
14	<0.36				0.46	39	5156.94	MgII_2796	0.8442
15	<0.36				0.55	39	5170.83	MgII_2803	0.8444
16	1.32	32	5267.46	CIV_1548	<0.31				
17	0.78	32	5276.16	CIV_1550	<0.30				
18					0.22	40	5319.04	...	...
19	0.30	34	5363.27	...					
20	<0.27				0.57	50	5488.50	FeII_2586	1.1219
21	<0.26				0.97	54	5516.90	FeII_2600	1.1217



### C.3 Spectra

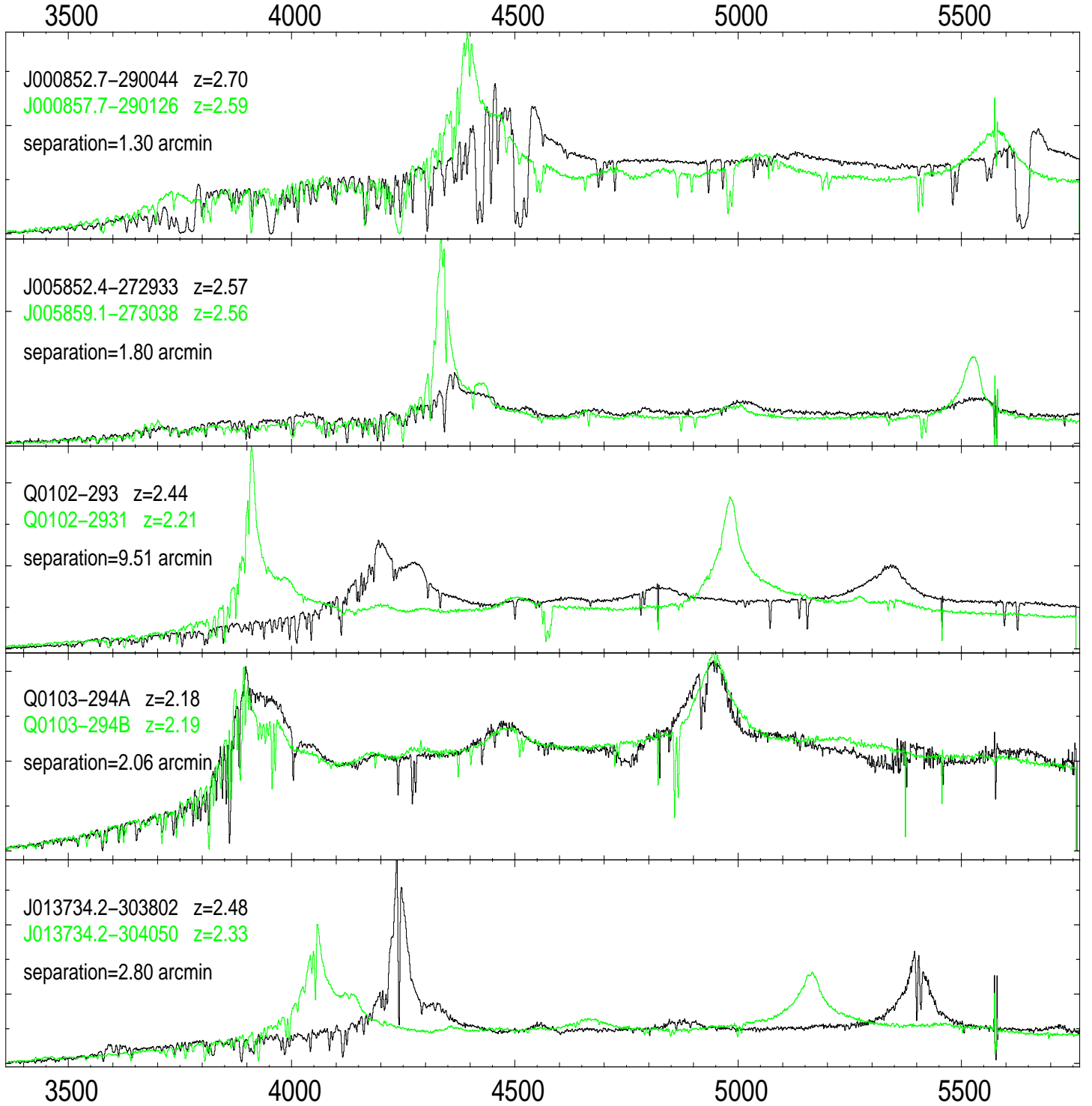


FIG. C.1 – Spectra of the observed QSO pairs in order of increasing right ascension. The names of the QSOs, their emission redshift and the separation between the two lines of sight are indicating in the top-left corner of each sub-panel.

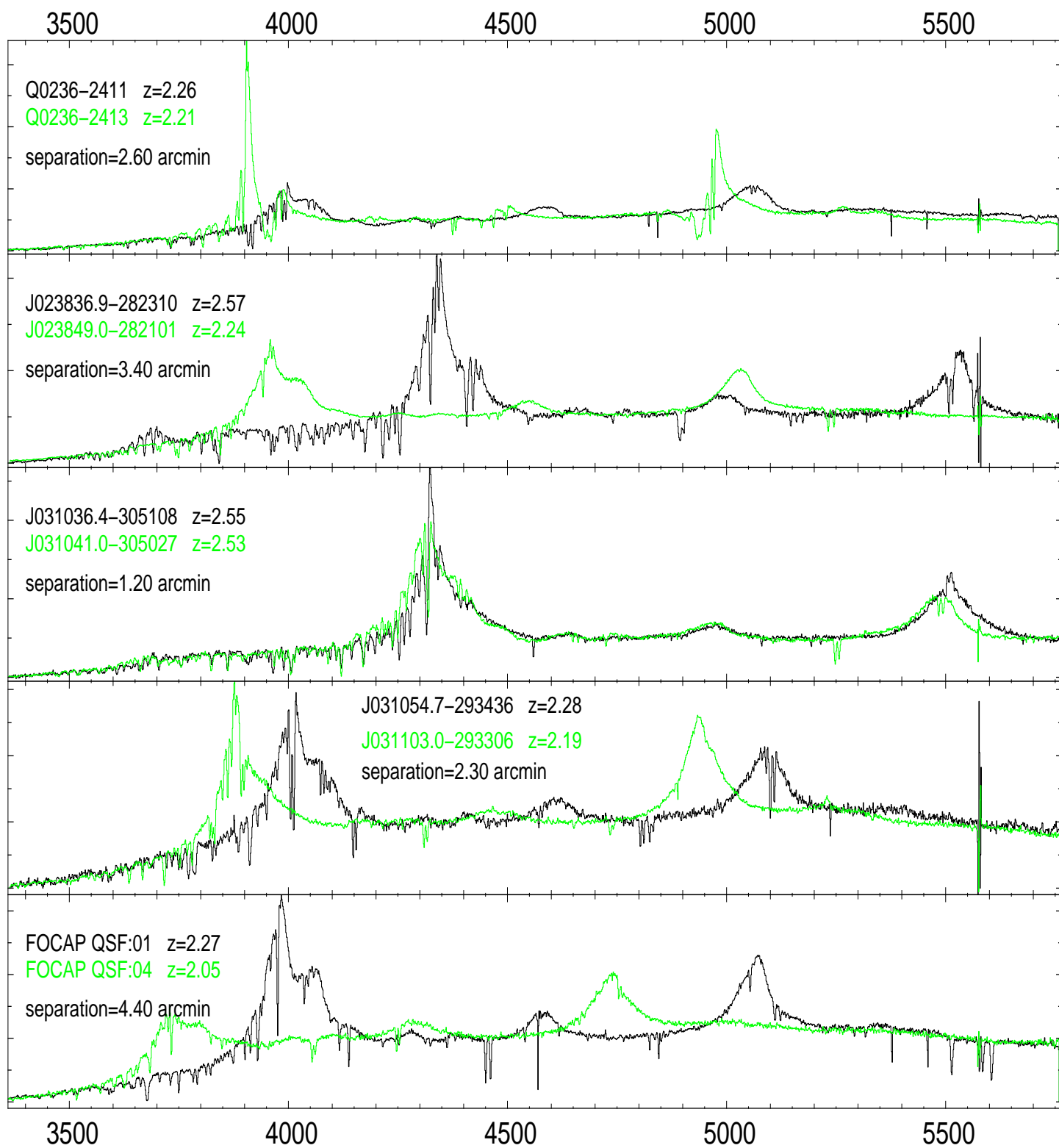


FIG. C.2 – As previously

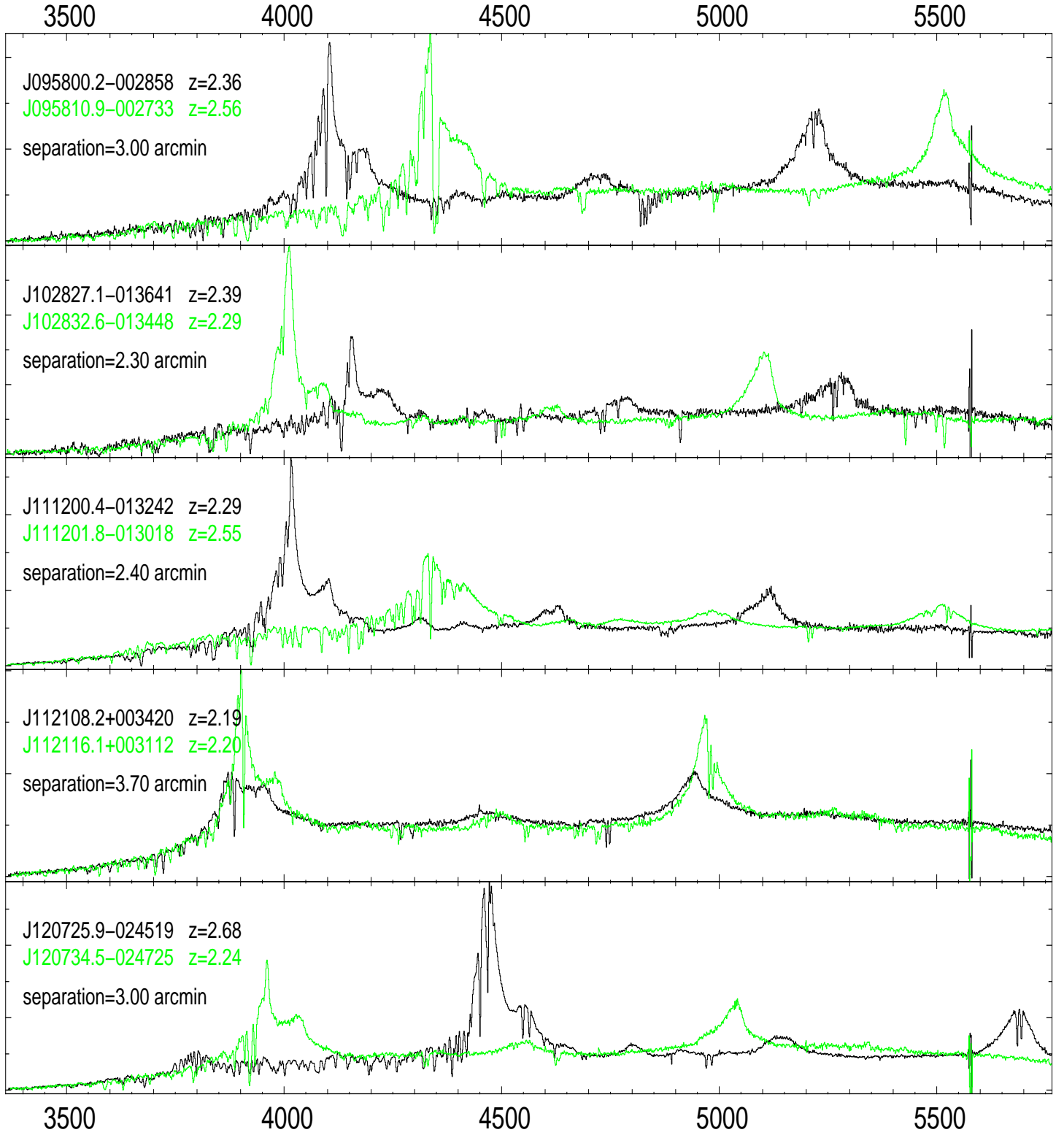


FIG. C.3 – As previously

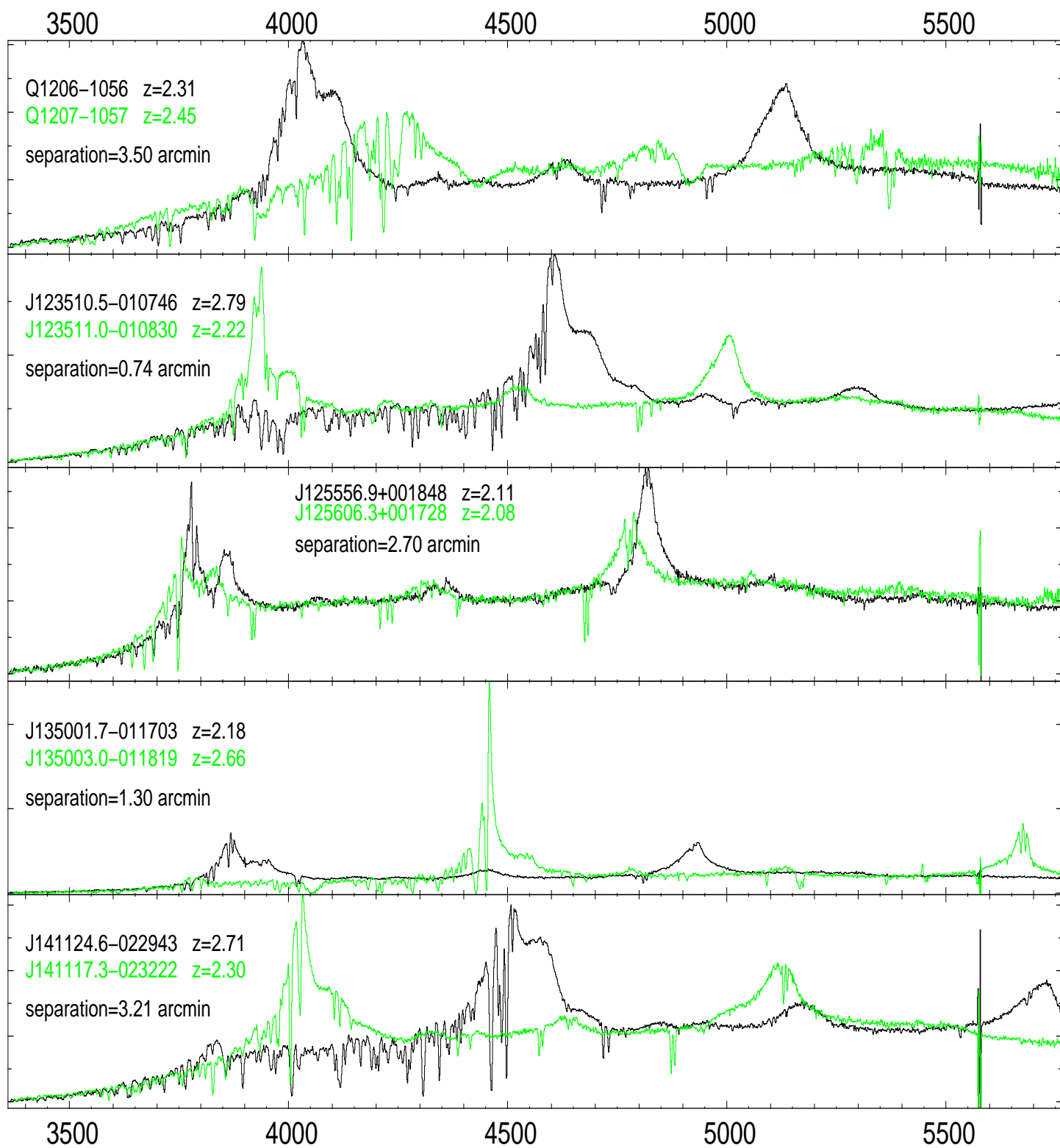


FIG. C.4 – As previously

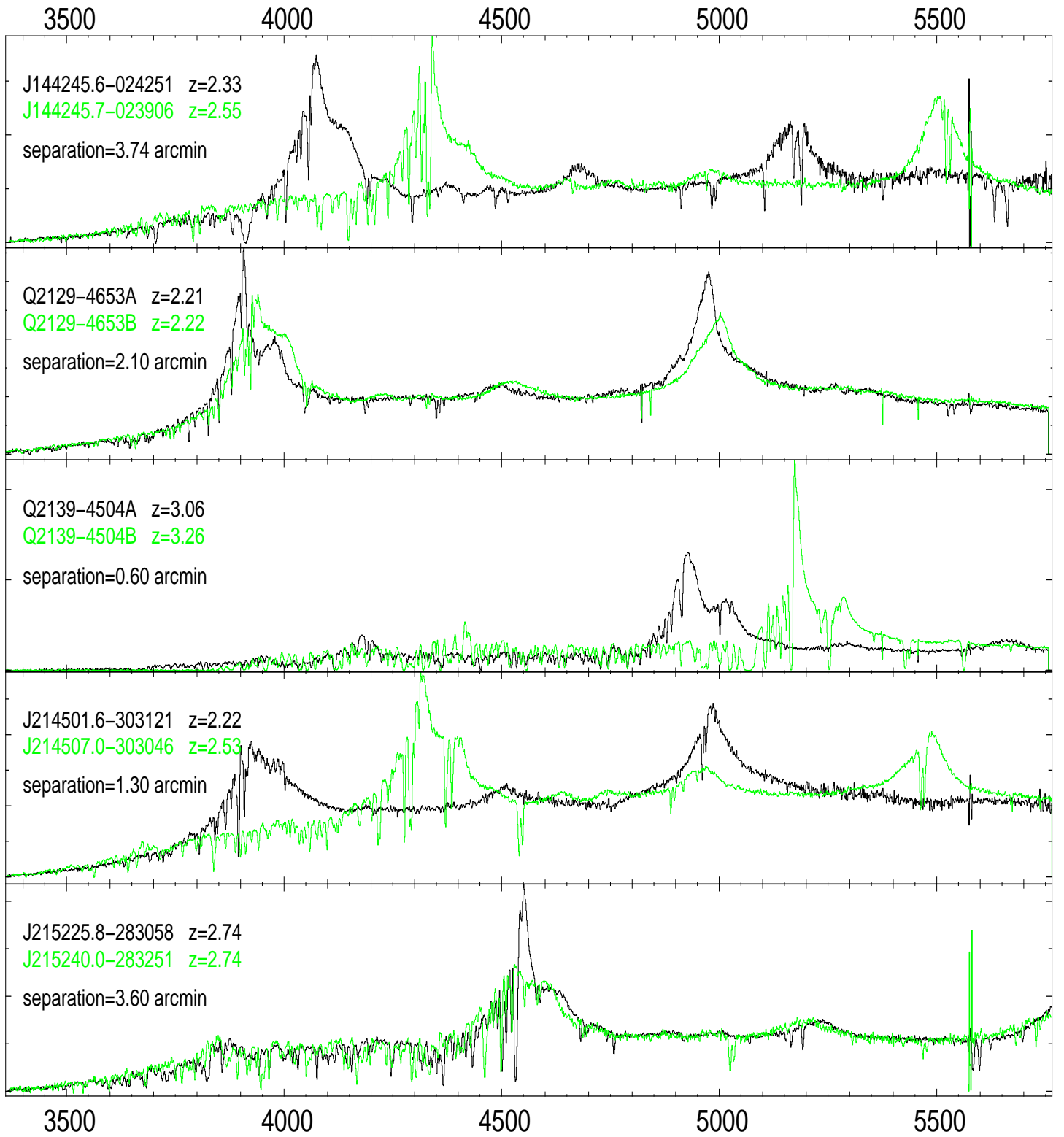


FIG. C.5 – As previously

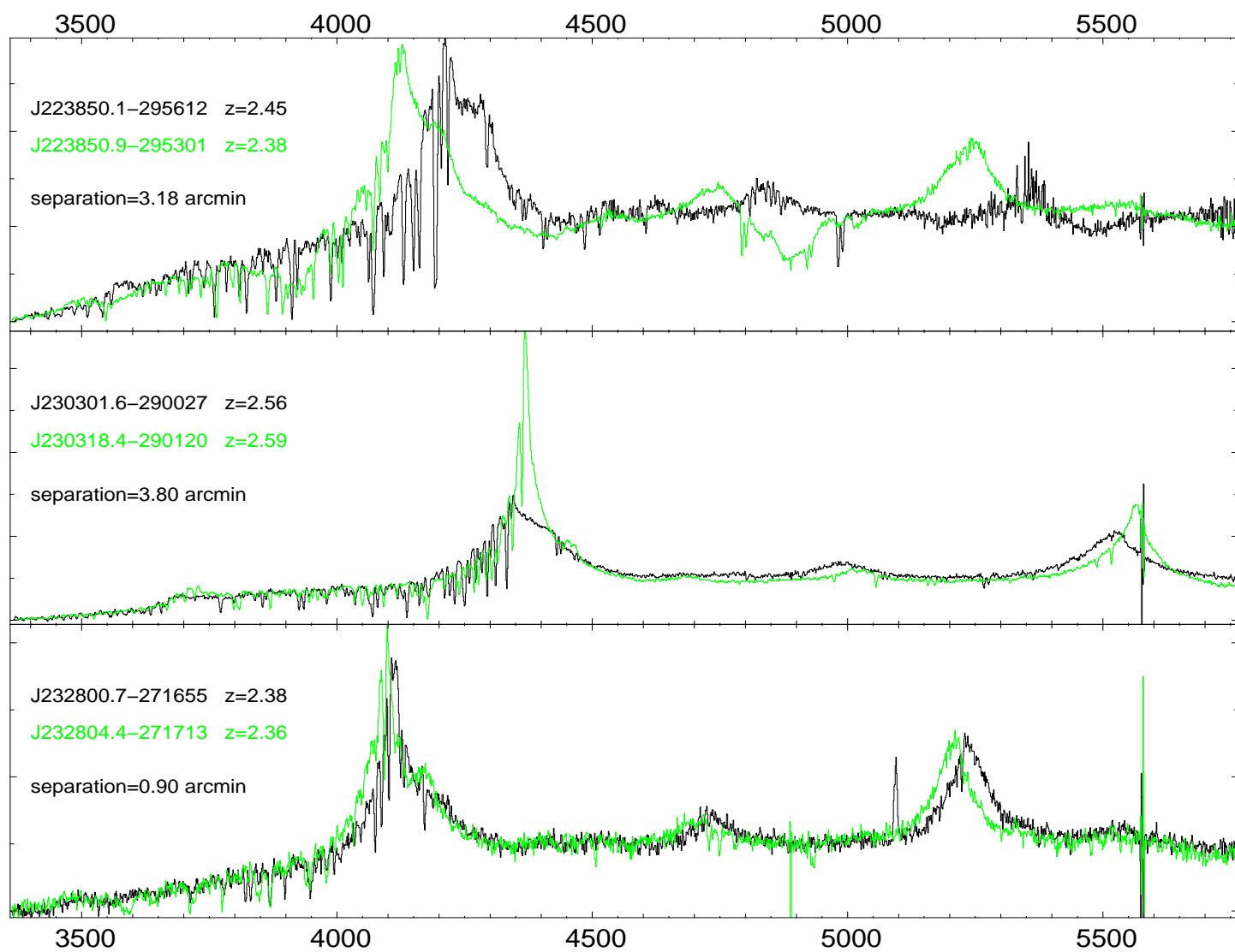


FIG. C.6 – As previously



## Annexe D

# HyperZ, ce qui a changé



## Version d'HyperZ

J'ai utilisé une version 'améliorée' de la version 1.1 d'HyperZ. Celle-ci m'a été fournie par Roser Pello. Elle comprend la possibilité d'un traitement a part pour les quasars, et surtout un programme 'classify-object' qui permet de calculer une probabilité normalisée à partir des différentes sorties issues d'HyperZ. J'ai effectué quelques changements directement dans le code fourni.

## Code principal

Plusieurs petites choses ont été modifiées : Au moment de l'écriture des fichiers de sortie. Afin d'obtenir une sortie du code homogène et ayant le même nombre de colonne (important pour la suite du traitement) dans le cas de non détection, j'ai modifié le code principal lignes 881 et 891 :

```
c ..... Write outputs in any case (except .m and .spe)
c ..... File .z_phot
zphot(1)=99.
r_chi2=99.
qq=0.
itypem=0
iagem=0
age=99.
av=0.
bb(1)=0.
DO i=1,3
  zz1(i)=99.
  zz2(i)=99.
END DO
z_poids=99.
qq_p=0.
amagb=99.
zphot(2)=99.
qq2=0.
amagb_clus=99.
p_clus=0.
amag(1)=99.
i_sol=1
iw_ob=0.
IF (z_clus.ge.0.) THEN
  IF (cz.eq.'y'.or.cz.eq.'Y') WRITE(17,2009) id,zphot(1),r_chi2,
  . qq,itypem,iagem,age,av,bb(1),zz1(1),zz2(1),zz1(2),
  . zz2(2),zz1(3),zz2(3),z_poids,qq_p,amagb,zphot(2),qq2,
  . amagb_clus,p_clus,amag(1)
ELSE
  IF (cz.eq.'y'.or.cz.eq.'Y') WRITE(17,2008) id,zphot(1),r_chi2,
  . qq,itypem,iagem,age,av,bb(1),zz1(1),zz2(1),zz1(2),
  . zz2(2),zz1(3),zz2(3),z_poids,qq_p,amagb,
  . zphot(2),qq2,iw_ob
END IF
```

Comme pour la version 1.2 je n'utilise plus les magnitudes limites dans le filtre de référence. Ce qui implique aussi une recherche de minimum dans un plus grand intervalle. Ainsi la taille des tableaux des solutions a été agrandie, ligne 60 :

```
DIMENSION izphot(532),zphot(532),bb(532),chi2sol(532)
```

## Code de la forêt Lyman-alpha

la version 1.2 d'HyperZ propose deux changements par rapport à celle que j'ai utilisé. D'une part une impossibilité d'avoir un facteur *da* négatif et d'autre part la possibilité de prendre en compte différentes absorptions (identique à Madau (1995), deux fois moindre et deux fois plus). J'ai repri le premier point, mais dans le second cas, j'ai préféré faire des tests et choisir lequel des trois cas représente le meilleur compromis.

```

SUBROUTINE lyman_forest(wl,fl,iw,z)
c ..... Computes the Lyman forest mean absorption of an
c input spectrum, according to Db and Da evolution
c from Madau (1995).
INTEGER mxw1,nw
PARAMETER (nw=100)
REAL fl,wl,da,db,z,zl,wtemp(nw),w1,w2,wstep
INCLUDE 'dimension_star.dec'
DIMENSION wl(mxw1),fl(mxw1),ptau(nw)
IF (z.lt.0.) THEN
zl=0.
ELSE
zl=z
END IF
w1=1050.(1.+zl)
w2=1170.(1.+zl)
wstep=(w2-w1)/FLOAT(nw)
DO i=1,nw
wtemp(i)=w1+(i-1)*wstep
ptau(i)=EXP(-3.6E-3*(wtemp(i)/1216.)3.46)
END DO
da=(1./(120.(1.+zl)))trapz1(wtemp,ptau,nw)
w1=920.(1.+zl)
w2=1015.(1.+zl)
wstep=(w2-w1)/FLOAT(nw)
DO i=1,nw
wtemp(i)=w1+(i-1)*wstep
ptau(i)=EXP(-1.7E-3*(wtemp(i)/1026.)3.46
-1.2E-3*(wtemp(i)/972.5)3.46
-9.3E-4*(wtemp(i)/950.)3.46)
END DO
db=(1./(95.(1.+zl)))trapz1(wtemp,ptau,nw)
More Lyman blanketing
da=0.5da
db=0.5db
c Less Lyman blanketing
c da=2.0da
c db=2.0db
c No Lyman blanketing
c da=1.0
c db=1.0
IF (da.gt.1.) da=1.
IF (db.gt.1.) db=1.
IF (da.lt.0.) da=0.
IF (db.lt.0.) db=0.
c correcting spectrum :
DO i=1,iw
IF (wl(i).lt.1216.0.and.wl(i).gt.1026.0) THEN!between Lya and Lyb
fl(i)=fl(i)da
END IF
IF (wl(i).le.1026.0) THEN!between Lyb and 912A break
fl(i)=fl(i)db
END IF
END DO
RETURN
END

```



## Annexe E

# 'The' selection, how to..

I will describe in this section how to do the selection on the technical point of view, on the deep fields of the CFHTLS. That it to say, this will precisely show what is needed, like all the programs written, to obtain a list of candidates in the deep fields of CFHTLS (as an example I will take D1). We will first see how to extract useful information from the material given by Terapix to produce a catalogue to use as an input of HyperZ. Then, this text will provide enough clues to use correctly HyperZ as it is needed to get the right outputs. These outputs will be required to make a first classification of the object, and then I will give the final selection via a Yorick code.

### E.1 Photometric catalogues

The data used from Terapix are the Catalogues-chi2, from Terapix : "These catalogues were produced using SExtractor in dual-image mode, using the gri chisquared image as reference, as well as the weight-map for the final stack. Each catalogue id for each filter for a given field refers to the same object. Note that these catalogues have not been corrected for galactic extinction. For almost all purposes, the Kron-like pseudo-total magnitudes in column 39 gives the best estimate of the object's magnitude (in non-crowded regions). Users who require z- or u- limited catalogues should use these files with caution. FILENAME : CFHTLS\_D\_[filter]chisq\_[RA][DEC]\_T0003.cat and CFHTLS\_[D/W]\_[ugriz/gri]\_chisq\_[RA][DEC]\_T0003.cat ." I use the non-multicolour catalogue that is to say the catalogue with names : CFHTLS\_D\_[filter]chisq\_[RA][DEC]\_T0003.cat. The description of the columns of this catalogues can be found in the photom-spica\_t03.param file, and you can see that for T0003 the Kron magnitude is not in colum 39 but 60. Once I have retrieved this files I use a awk code to extract some informations :

```

awk 'print($20,$21,$60,$61,$80,$81)' CFHTLS_chisq_gri_022559-042940_u_T0003.cat >
D1_u_refchi2_selection.dat
awk 'print($60,$61,$80,$81)' CFHTLS_chisq_gri_022559-042940_g_T0003.cat >
D1_g_refchi2_selection.dat
awk 'print($60,$61,$80,$81)' CFHTLS_chisq_gri_022559-042940_r_T0003.cat >
D1_r_refchi2_selection.dat
awk 'print($60,$61,$80,$81)' CFHTLS_chisq_gri_022559-042940_i_T0003.cat >
D1_i_refchi2_selection.dat
awk 'print($60,$61,$80,$81)' CFHTLS_chisq_gri_022559-042940_z_T0003.cat >
D1_z_refchi2_selection.dat

```

I take from these catalogues, the position from  $u$ ,  $ra$  and  $dec$  : \$20,\$21. The magnitude and its error : \$60,\$61. And finally the *CLASS\_STAR* flag and the quality flag : \$80,\$81. From this catalogues I will make only one and select the objects with  $17 < i < 23$ , with at least one of the error in one of the filter below 0.2, with the quality flags below or equal to 3 (you can find the signification of these flags in the Source Extractor Manual) and with a *CLASS\_STAR* above 0.5.

```

paste D1_u_refchi2_selection.dat D1_g_refchi2_selection.dat D1_r_refchi2_selection.dat
D1_i_refchi2_selection.dat D1_z_refchi2_selection.dat > cat_tmp.dat
awk 'if($15<23 && $15>17 && ($4 <=0.2 || $8 <=0.2 || $12<=0.2 || $16<=0.2 ||
$20<=0.2) && $5<=3 && $9<=3 && $13<=3 && $17<=3 && $21<=3 && $18 >= 0.5)
print($1,$2,$3,$7,$11,$15,$19,$4,$8,$12,$16,$20)' cat_tmp.dat > D1_ap_cat_coord_chi2.dat

```

Then I will remove all the objects that are inside the mask regions. From Terapix : "Built using the "polygon" option of the DS9. visualisation tool to create a "\*.reg" file which describes, in world co-ordinates, the regions of the images which should not be used for scientific analysis. Can be used with DS9 (or SAOimage, after they have been transformed to pixel co-ordinates). They mask the MEGACAM field boundary, saturated stars, satellite tracks and image defects, if any. It is an ASCII file : polygons may be added or removed according to user wishes It is important that any users check the masks and modify them according to its scientific needs. Note that the masks are constructed using the I-band images which are the references for all filters. so they can be used with the u,g,r and z images. Users who require catalogues selected in other bands (for instance, u or z) many need to make some changes to these masks. FILENAME : CFHTLS\_[D/W]\_[filter]\_[RA][DEC]\_T0003.reg." In fact there is only one mask file taken in the  $i$  band. Moreover sometimes the mask files are not in world coordinates, but in pixels. In that case you should load the file in pixel in DS9 with an image of the field (preferably multi-colour or  $i$  band) and save the region in world units. In the aim to remove the objects I use a program belongpol.c (source code in figure E.1 and E.2) which need in entry the .reg file provided by Terapix and the catalogue created before :

```

belongpol.exe CFHTLS_D_i_022559-042940_T0003.reg D1_ap_cat_coord_chi2.dat
D1_ap_cat_coord_mask_chi2.dat

```

Then I create a file .reg which can be used by DS9 to plot the catalogue of sources. It is really useful to check either you made the right choice for the flag and for the

*CLASS\_STAR* parameters since you can visualise the objects (better to use with the multi-colour image) :

```
awk 'print("ICRS;ellipse("$1","$2","0.0001","0.0001",45))' D1_ap_cat_coord_mask_chi2.dat
> temp.reg
paste deb.reg temp.reg > D1_mask.reg
```

Note that you can change the 0.0001 by a factor of the magnitude of the object. With deb.reg, a header of a classical .reg regions file :

```
global color=red font="helvetica 10 normal" select=1 edit=1 move=1 delete=1 include=1 fixed=0
source
```

Then I create a file which could be used as an entry for HyperZ, but still need to be corrected for the dereddening and to account for the calibration errors :

```
awk 'print(NR,$3,$4,$5,$6,$7,$8,$9,$10,$11,$12)' D1_ap_cat_coord_mask_chi2.dat >
D1_ap_cat_HZ_mask_chi2.dat
rm -f temp.reg cat_tmp.dat D1_u_refchi2_selection.dat D1_g_refchi2_selection.dat
D1_r_refchi2_selection.dat D1_i_refchi2_selection.dat D1_z_refchi2_selection.dat
```

So finally I have to apply the dereddening values (all you need to know about how it works is there : <http://astro.berkeley.edu/davis/dust/data/data.html> and the use of wcs tools (for skycoor) : <http://tdc-www.harvard.edu/wcstools/> ) :

```
awk 'print($1,$2,"J2000")' D1_ap_cat_coord_mask_chi2.dat > tt.dat
skycoor -g @tt.dat > gal.dat
dust_getval infile=gal.dat interp=y ipath=/DATA/Dered outfile=dered.dat
awk 'print(5.155*$3,3.793*$3,2.751*$3,2.086*$3,1.479*$3)' dered.dat > tt2.dat
paste D1_ap_cat_coord_mask_chi2.dat tt2.dat > Cat_out_dered_temp.dat
```

here I will add the errors due to the dereddening and to the calibration that are estimated for the former to 15% of the dereddening value and of 0.05 for the latter. Then I transform the files in an input for HyperZ

```
awk 'print($1,$2,$3-$13,$4-$14,$5-$15,$6-$16,$7-$17, sqrt($8*$8+0.0025+0.02255*$13*$13),
sqrt($9*$9+ 0.0025+ 0.02255*$14*$14), sqrt($10*$10+0.0025+ 0.0225*$15*$15), sqrt($11*$11+
0.0025+ 0.0225*$16*$16), sqrt($12*$12+0.0025+ 0.0225*$17*$17))' Cat_out_dered_temp.dat
> Cat_outdered_temp2.dat
awk 'print(NR,$0)' Cat_outdered_temp2.dat > Cat_outdered_coord_mask.dat
awk 'print(NR,$3,$4,$5,$6,$7,$8,$9,$10,$11,$12)' Cat_outdered_temp2.dat >
Cat_outdered_HZ_mask.dat
rm -f tt.dat tt2.dat gal.dat dered.dat Cat_out_dered_temp.dat Cat_outdered_temp2.dat
```

At the end of this process the file Cat\_outdered\_HZ\_mask.dat is ready to be used as an input for HyperZ and the file Cat\_outdered\_coord\_mask.dat will be useful to crosscorrelate or to extract the position of the candidates at the end of the selection. It is possible to check the catalogue by looking at the D1\_mask.reg in a DS9 window with

224	0.00	1	27.3	0.5
225	0.00	1	27.4	0.5
226	0.00	1	27.4	0.5
227	0.00	1	27.45	0.5
228	0.00	1	27.51	0.5

TAB. E.1 – content of the filter parameter

an image of the field. Note something very important here. I made the choice to select my sources before running hyperz, so it takes less time to run it. But if you want to change the limiting magnitudes, the flags or other things in your input files you will have to run again hyperz. So there is another way to do it : you run hyperz on all the sources contained in the catalogue, without any constraints on the flag or magnitude, then you run hyperz, and finally you select your objects at the same time than your candidates. It will take some day for hyperz to run, and you will have to manipulate for a long time large files, but it is a good choice if you know by advance that you will have to make a lot of changes but to work with the same catalogues (so I will advice to use this only with the final distribution of Terapix catalogues).

## E.2 How to use HyperZ

We need to run hyperZ three times to get the right outputs, one for each kind of objects : galaxies, stars and quasars. First the program called simply 'hyperz', which is the last version of hyperz. These one will give us the output for galaxies, we don't really need it since the data here will not be used in our selection. The parameter file is given figure E.3. Then we use the hyper\_star program, which has been modify accordingly to the choice of the treatment for the Lyman- $\alpha$  forest in hyper\_star\_ply or hyper\_star\_mly, for more or less lyman blanketing. This program is run using first the star.param, which will compute  $\chi^2$  for star templates and then with qso.param which will compute them for quasar templates. These two files can be seen in figure E.4 and E.5 respectively. The files containing the list of templates used for quasars and stars can be found in appendix F The filter files for CFHT is described table E.1. Don't forget that the response filter (which are referenced here from 224 to 228 need to take into account response from the filter and the detector).

Once all the outputs are given, we use them to provide another output which allow us to do a first classification using the script untitled 'Classify\_D1'. This script provide a file called D1\_T0003.clas, containing the classification of the objects according to the program 'clasify\_object'. In this program (directly inside the code source) you will have to choose what are the thresholds for the probability, I choose  $p(1)=0.05$  and  $p(2)=0.99$  (see the code for further explanations).

## E.3 Final selection

The final selection will use all these outputs to discriminate the quasars candidates. This is done with a yorick code (note that the previous pre-selection could be translated

into this yorick code). I will give here (figure E.6) only the different selections available in the code, the signification can be found in the chapter 4. It is possible to mix the selection, that is to say : the loose selection can be used until a limit magnitude still yet to determine, probably around 20.3 in  $g$  and then use the optimal selection to limit the number of candidates above a given magnitude.

The code provides finally a list of candidates in an ascii format with the coordinates and the photometric redshift, as well as a .reg file in order to plot them on a DS9 window, so you can cross-check visually with a .reg file containing the quasars of the field.



FIG. E.1 – belongpol.c - 1

25 jui 06 13:48

belongpol.c

Page 1/3

```

/*****
Written By Franck 25/12/2005

english :

Test either a point belongs to a polygon
Input 1 is a file with the corners of the pygons, like a .reg DS9 file.
And Input 2 is a file with the coordinates in the first 2 columns

The output copy the input file containing the right points.

French:

Test si un point appartient a un polynome,
en entree un fichier contenant les points des coins des polynome
de type .reg de ds9
un fichier de coordonee
et le fichier de sortie qui recopie le fichier de coordonnees.

fonction de lecture du fichier et ecriture resultat
fonction de test d'appartennace a un polygone

*****/

#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include <math.h>

int pointInPolygon(int polySides,float *polyX,float *polyY,float x,float y);
int readregions(char *nom,int polySides,float x,float y);

int main(int argc, char *argv[])
{
    int polySides,result;
    char *buffer1,*buffer2;
    char *part1,*part2;
    float x;
    float coord_px;
    float y;
    float coord_py;

    FILE *coord=fopen(argv[2],"r");
    if(coord==NULL){
        fprintf(stderr,"Erreur dans l'ouverture du fichier");
        exit(-1);
    }
    FILE *output=fopen(argv[3],"w");
    buffer1=(char*) malloc(500);
    buffer2=(char*) malloc(500);
    while(fgets(buffer1,500,coord))
    {
        sprintf(buffer2,buffer1);
        part1=strtok(buffer2," ");
        part2=strtok(NULL," ");
    }
}
```

25 jui 06 13:48

belongpol.c

Page 2/3

```

    sscanf(part1,"%f",&coord_px);
    sscanf(part2,"%f",&coord_py);
    result= readregions(argv[1],polySides,coord_px,coord_py);
    if(result==0) fprintf(output,buffer1);
}
free(buffer1);
free(buffer2);
fclose(coord);
fclose(output);
return 0;
}

int readregions(char *nom, int polySides, float x, float y)
{
    char* buffer;
    char *part1,*part2,*part3,*part4;
    char *lecture;
    int i;
    float coordX,coordY;
    int belong = 0;
    float polyX[100];
    float polyY[100];
    FILE *regions=fopen(nom,"r");
    if(regions==NULL){
        fprintf(stderr,"Erreur dans l'ouverture du fichier");
        exit(-1);
    }

    buffer=(char*) malloc(500);

    fgets(buffer,500,regions);
    fgets(buffer,500,regions);
    fgets(buffer,500,regions);
    part1=(char*) malloc(500);
    part2=(char*) malloc(500);
    part3=(char*) malloc(500);
    part4=(char*) malloc(500);

    while(belong==0)
    {
        lecture = fgets(buffer,500,regions);
        if(lecture!=NULL)
        {
            part1=strtok(buffer,"(");
            part2=strtok(NULL,")");
            part3=part2;
            i=0;
            while(part3!=NULL)
            {
                if(i==0) part3=strtok(part2,",");
                else part3=strtok(NULL,",");
                if(part3!=NULL)
                {
                    part4=strtok(NULL,"");
                    sscanf(part3,"%f",&coordX);
                    sscanf(part4,"%f",&coordY);
                    polyX[i]=coordX;
                    polyY[i]=coordY;
                }
            }
        }
    }
}
```

Fig. E.2 – belongpol.c - 2

25 jui 06 13:48	<b>belongpol.c</b>	Page 3/3
-----------------	--------------------	----------

```

        i++;
    }
    else polySides = i;
}
// a partir d'ici j'ai les coins d'un polygone, test si le point est d
edans...
    belong=pointInPolygon(polySides,polyX,polyY,x,y);
}
else
{
    free(buffer);
    fclose(regions);
    return belong;
}
}
free(buffer);
fclose(regions);
return belong;
}

int pointInPolygon(int polySides,float *polyX,float *polyY,float x,float y)
{
    int i,j=0;
    int oddNODES=0;
    for (i=0;i<polySides;i++)
    {
        j++;
        if (j==polySides) j=0;
        if (polyY[i]<y && polyY[j]>=y || polyY[j]<y && polyY[i]>=y)
        {
            if(polyX[i]*(y-polyY[j])-(polyY[j]-polyY[i])*(polyX[j]-polyX[i])<x)
            {
                if(oddNODES==0) oddNODES=1;
                else oddNODES=0;
            }
        }
    }
    return oddNODES;
}

```

FIG. E.3 – hyperz\_galaxies.param

07 aoÅ» 06 12:43	hyperz_CFHT.param	Page 1/2
#####	hyperz & make_catalog parameters #####	
#####	version 1.2 - October 31, 2001 #####	
#####	by Bolzonella, Pello' & Miralles #####	
#####	#####	
####	Reference: Bolzonella, Miralles & Pello', 2000, A&A 363, 476 ####	
####	Web: http://webast.ast.obs-mip.fr/hyperz/ ####	
#####	#####	
A0VSED	/home/ogre/HyperZ/HyperZ_New/ZPHOT1.3_galaxie/filters/A0V_KUR_BB.SE	
D	# Vega SED	
FILTERS_RES	/home/ogre/HyperZ/HyperZ_New/ZPHOT1.3_galaxie/filters/FILTER.RES	
	# filters' transmission	
FILTERS_FILE	filters/CFHT.param # filter's file	
#FILTERS_FILE	filter_SDSS.param # filter's file pour cat	
TEMPLATES_FILE	spectra.param # templates	
#TEMPLATES_FILE	/home/ogre/HyperZ/HyperZ_New/qso_star/qso_spectra2 # templat	
es	pour catalogue de QSO	
s2	# templates pour catalogue de star	
MAG_TYPE	1 # 0: standard Vega magnitudes	
	# 1: AB magnitudes (Oke 1974)	
ERR_MAG_MIN	0.01 # minimum photometric error	
	# (used in make_catalog if NOISE_TYPE=1)	
Z_MIN	0. # minimum redshift	
Z_MAX	1. # maximum redshift	
REDDENING_LAW	0 # reddening law	
	# 0: no reddening	
	# 1: Allen (1976) MW	
	# 2: Seaton (1979) MW	
	# 3: Fitzpatrick (1983) LMC	
	# 4: Prevot (1984) Bouchet (1985) SMC	
	# 5: Calzetti (2000)	
AV_MIN	0.0 # minimum A_v	
AV_MAX	1.5 # maximum A_v	
LY_FOREST	1.0 # Lyman Forest	
	# 1. : fixed Madau value	
	# other values: 3 optical depth considered	
	# <tau_eff>/LY_FOREST, <tau_eff> and	
	# <tau_eff>*LY_FOREST	
#OUTPUT_FILE	models/star_0_1 # cat star	
#OUTPUT_FILE	models/qso_0_6 # N.B. no extension !! pour cat	
OUTPUT_FILE	D1.Galaxies # N.B. no extension !!	
	# hyperz: root name for output files	
	# (see other parameters)	
	# make_catalog: .model random catalogue	
	# .cat catalogue with noise	
H0	70. # Hubble constant in Km/s/Mpc	
OMEGA_M	0.7 # density parameter (matter)	
OMEGA_V	0.3 # density parameter (Lambda)	
#####	hyperz only #####	
FILT_M_ABS	91 # filter for absolute magnitude	
CATALOG_FILE	/home/ogre/HyperZ/HyperZ_New/Catalogues/D1_Cat_outdered_HZ_mask.	
dat		
CATALOG_TYPE	0 # catalogue type	
	# 0: z/cat	
	# 1: z/obj	
Z_STEP	0.2 # step in redshift	
ZSTEP_TYPE	0 # 0 step = Z_STEP	
	# 1 step = Z_STEP*(1+z)	
AV_STEP	0.30 # Av_err	

lundi 07 aoÅ»t 2006

hyperz\_CFHT.param

07 aoÅ» 06 12:43	hyperz_CFHT.param	Page 2/2
AGE_CHECK	y # check age gal. < age universe	
PROB_THRESH	10.00 # prob. thresh. for second. max. (0,100)	
OUTPUT_TYPE	0 # 0: 2.E-17 erg/cm^2/s/A,	
	# 1: microJy, 2: mag_AB	
SED_OBS_FILE	n # file .obs_sed	
SED_TEMP_FILE	n # file .temp_sed	
LOGPHOT_FILE	n # file .log_phot	
CATPHOT_FILE	n # file .cat_phot	
ZPHOT_FILE	y # file .z_phot	
#####	optional parameters #####	
#Z_CLUSTER	1. # redshift of cluster	
#M_ABS_MIN	-28. # minimum absolute magnitude (bright)	
#M_ABS_MAX	-9. # maximum absolute magnitude (faint)	
#MATRIX	n # file .m for each object	
#SPECTRUM	n # file .spe for each object	
#EBV_MW	0. # E(B-V) for galactic dereddening	
#####	make_catalog only #####	
REF_FILTER	113 # reference filter	
MAG_MIN	22. # minimum apparent magnitude in REF_FILTER	
MAG_MAX	25. # maximum apparent magnitude in REF_FILTER	
NOISE_TYPE	0 # 0: random noise with fixed sigma in mag	
	# 1: noise dm(m)	
N_OBJ	1000 # number of objects	
Z_FORM	10. # formation redshift	
AGE_RANDOM	0 # 0 age = age of universe - T(z_form)	
	# 1 age = random in [0,T(z)-T(z_form)]	
RAND_SEED	3010 # seed for random generator (max 4 figures)	
#####	#####	

1/1

FIG. E.4 – hyperz\_stars.param

07 août 06 12:47	toto.txt	Page 1/1
# Parameters file		
AOVSED	/home/ogre/HyperZ/HyperZ_New/ZPHOT1.3_galaxie/filters/AOV_KUR_BB.SE	
D	# Vega SED	
FILTERS_RES	/home/ogre/HyperZ/HyperZ_New/ZPHOT1.3_galaxie/filters/FILTER.RES	#
	filters' transmission	
FILTERS_FILE	/home/ogre/HyperZ/HyperZ_New/ZPHOT1.3_galaxie/filters/CFHT.param	
	# filters' file	
TEMPLATES_FILE	star_spectra_WD	# models file
FILT_M_ABS	91	# filter for absolute magnitude
ERR_MAG_MIN	0.0001	# err_min
Z_MIN	-0.002	# minimum redshift
Z_MAX	0.002	# maximum redshift
Z_STEP	0.0001	# step in redshift
ZSTEP_TYPE	0	# 0 step = Z_STEP
	1	# 1 step = Z_STEP*(1+z)
CATALOG_TYPE	0	# cat.type
	0	# 0 z/cat
	1	# 1 z/obj
CATALOG_FILE	/home/ogre/HyperZ/HyperZ_New/Catalogues/D1_Cat_outdered_HZ_mask.dat	
MAG_TYPE	1	# 0 standard Vega mag, 1 AB mag
REDDENING_LAW	0	# reddening law
	0	# 0 no reddening
	1	# 1 Allen (1976) MW
	2	# 2 Seaton (1979) MW
	3	# 3 Fitzpatrick (1983) LMC
	4	# 4 Prevot (1984) Bouchet (1985) SMC
	5	# 5 Calzetti (1999)
AV_MIN	0.00	# Av_min
AV_MAX	1.50	# Av_max
AV_STEP	0.30	# Av_err
AGE_CHECK	n	# check age gal. < age universe
PROB_THRESH	10.00	# prob. thresh. for second. max. (0,100)
OUTPUT_FILE	D1.stars	# name of output files (no extension)
OUTPUT_TYPE	0	# 0 2.E-17 erg/cm^2/s/A, 1 microJy, 2 mag_AB
SED_OBS_FILE	n	# file .obs_sedq
SED_TEMP_FILE	n	# file .temp_sed
LOGPHOT_FILE	n	# file .log_phot
CATPHOT_FILE	n	# file .cat_phot
ZPHOT_FILE	y	# file .z_phot
H0	70	# Hubble constant
		# used to compute absolute magnitude and
		# ages if AGE_CHECK = y
OMEGA_M	0.3	# density parameter (matter)
		# used to compute absolute magnitude and
		# ages if AGE_CHECK = y
OMEGA_V	0.7	# density parameter (Lambda)18.51
		# used to compute absolute magnitude and
		# ages if AGE_CHECK = y
#####	cluster option	#####
#Z_CLUSTER	0.02	# redshift of cluster
#####	optional parameters	#####
#M_ABS_MIN	-27.	# minimum absolute magnitude (bright)
#M_ABS_MAX	-5.	# maximum absolute magnitude (faint)
MATRIX	n	# file .m for each object
SPECTRUM	n	# file .spe for each object
EBV_MW	0.	# E(B-V) for galactic dereddening



Fig. E.6 – The different selections available

07 août 2006 15:04	toto.i	Page 1/1
<pre>//== selection optimal SDSS // selection = where((fres(,4)!=0 &amp; fres(,3)==0 &amp; fres(,5) &lt;= 5. &amp; output(,3) &lt;= 60. &amp; str2float(field(,5)) &lt;= lim_mag &amp; fres(,5) &lt;= limz_phot_sup &amp; fres(,5) &gt;= limz_phot_inf)   ( fres(,5) &lt;= 5. &amp; output(,3) &lt;= 1. &amp; fres(,6) &gt; 3. &amp; str2 float(field(,5)) &lt;= lim_mag &amp; fres(,5) &gt;= 2.6 &amp; fres(,5) &lt;= limz_phot_sup &amp; fres (,5) &gt;= limz_phot_inf ));  //==== selection large //selection = where( (fres(,4)!=0 &amp; fres(,3)==0 &amp; fres(,5) &lt;= 5. &amp; output(,3) &lt;= 60. &amp; str2float(field(,5)) &lt;= lim_mag &amp; fres(,5) &lt;= limz_phot_sup &amp; fres(,5) &gt;= limz_phot_inf)   (output(,3) &lt;= fres(,6) &amp; str2float(field(,5)) &lt;= lim_mag &amp; output(,3)&lt;9999.)   ( fres(,5) &lt;= 5. &amp; output(,3) &lt;= 1. &amp; fres(,6) &gt; 3. &amp; str2 float(field(,5)) &lt;= lim_mag &amp; fres(,5) &gt;= 2.6 &amp; fres(,5) &lt;= limz_phot_sup &amp; fres (,5) &gt;= limz_phot_inf));  //==== selection to avoid the redshift interval 2.2 - 2.35 --- Loose Selectio n // selection = where((output(,3) &lt; fres(,6) &amp; str2float(field(,5)) &lt;= lim_ma g &amp; str2float(field(,7)) &gt; lim_magi &amp; fres(,5) &lt;= 5. &amp; (fres(,5) &lt; 2.2   fres(, 5) &gt; 2.4) &amp; fres(,5) &lt;= limz_phot_sup &amp; fres(,5) &gt;= limz_phot_inf)   (output(,3) &lt; fres(,6) &amp; str2float(field(,5)) &lt;= lim_mag &amp; str2float(field(,7)) &gt; lim_magi &amp; fres(,5) &lt;= 5 &amp; output(,3)/max(fres(,6),10E-15) &lt; 0.3 &amp; (fres(,5) &gt;= 2.2   f res(,5) &lt;= 2.4) &amp; fres(,5) &lt;= limz_phot_sup &amp; fres(,5) &gt;= limz_phot_inf));  //==== selection to avoid the redshift interval 2.2 - 2.35 and 0-0.13 --- Opt imal Selection selection = where((output(,3) &lt; fres(,6) &amp; str2float(field(,5)) &lt;= lim_ma g &amp; str2float(field(,7)) &gt; lim_magi &amp; fres(,5) &lt;= 5. &amp; (fres(,5) &lt; 2.2   fres(, 5) &gt; 2.4) &amp; fres(,5) &gt; 0.25 &amp; fres(,5) &lt;= limz_phot_sup &amp; fres(,5) &gt;= limz_phot_ inf)   (output(,3) &lt; fres(,6) &amp; str2float(field(,5)) &lt;= lim_ma g &amp; str2float(field(,7)) &gt; lim_magi &amp; fres(,5) &lt;= 5 &amp; output(,3)/max(fres(,6),1 0E-15) &lt; 0.3 &amp; (fres(,5) &gt;= 2.2 &amp; fres(,5) &lt;= 2.4) &amp; fres(,5) &lt;= limz_phot_sup &amp; fres(,5) &gt;= limz_phot_inf)   (output(,3) &lt; fres(,6) &amp; str2float(field(,5)) &lt;= lim_mag &amp; st r2float(field(,7)) &gt; lim_magi &amp; fres(,5) &lt;= 5 &amp; output(,3)/max(fres(,6),10E-15) &lt; 0.1 &amp; (fres(,5) &gt;= 0.0 &amp; fres(,5) &lt;= 0.25) &amp; fres(,5) &lt;= limz_phot_sup &amp; fres(, 5) &gt;= limz_phot_inf));  //==== Basic selection // selection = where(output(,3) &lt; fres(,6) &amp; str2float(field(,5)) &lt;= lim_mag &amp; fres(,5) &lt;= 5. &amp; fres(,5) &lt;= limz_phot_sup &amp; fres(,5) &gt;= limz_phot_inf);  //==== selection without chi2 value //selection = where( (fres(,4)!=0 &amp; fres(,3)==0 &amp; fres(,5) &lt;= 5. &amp; str2float( field(,5)) &lt;= lim_mag &amp; fres(,5) &lt;= limz_phot_sup &amp; fres(,5) &gt;= limz_phot_inf)   ( output(,3) &lt;= fres(,6) &amp; str2float(field(,5)) &lt;= lim_mag ));</pre>		



## Annexe F

# List of templates

List of star templates used in appendix E :

[illegible]



List of quasars templates (from spectroscopic selection) used in appendix E :

```

/home/ogre/HyperZ/HyperZ_New/qso_star/qso_templates2/1.spe A
/home/ogre/HyperZ/HyperZ_New/qso_star/qso_templates2/10.spe A
/home/ogre/HyperZ/HyperZ_New/qso_star/qso_templates2/11.spe A
/home/ogre/HyperZ/HyperZ_New/qso_star/qso_templates2/13.spe A
/home/ogre/HyperZ/HyperZ_New/qso_star/qso_templates2/14.spe A
/home/ogre/HyperZ/HyperZ_New/qso_star/qso_templates2/15.spe A
/home/ogre/HyperZ/HyperZ_New/qso_star/qso_templates2/17.spe A

```

/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/18.spe A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/19.spe A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/2.spe A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/3.spe A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/4.spe A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/5.spe A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/6.spe A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/7.spe A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/8.spe A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/9.spe A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/QSO\_11.dat A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/QSO\_21.dat A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/QSO\_25.dat A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/QSO\_31.dat A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/HST\_MODIF\_14\_03.dat A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/HST\_MODIF\_14\_044.dat A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/HST\_MODIF\_14\_075.dat A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/HST\_MODIF\_176\_075.dat A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/HST\_MODIF\_22\_044.dat A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/HST\_MODIF\_24\_044.dat A  
/home/ogre/HyperZ/HyperZ\_New/qso\_star/qso\_templates2/hstdsslow2.dat A



# Bibliographie

- Abazajian, K., Adelman-McCarthy, J. K., Agüeros, M. A., & et al. 2004, *AJ*, **128**, 502
- Adelberger, K. L., Steidel, C. C., Shapley, A. E., & Pettini, M. 2003, *ApJ*, **584**, 45
- Aguirre, A., Schaye, J., & Theuns, T. 2002, *ApJ*, **576**, 1
- Alcock, C. & Paczynski, B. 1979, *Nature*, **281**, 358
- Aracil, B., Petitjean, P., Pichon, C., & Bergeron, J. 2004, *A&A*, **419**, 811
- Aracil, B., Petitjean, P., Smette, A., Surdej, J., Mückel, J. P., & Cristiani, S. 2002, *A&A*, **391**, 1
- Aracil, e. a. 2006, *in preparation*
- Arons, J. & Wingert, D. W. 1972, *ApJ*, **177**, 1
- Babbedge, T. S. R., Rowan-Robinson, M., Gonzalez-Solares, E., Polletta, M., Berta, S., Pérez-Fournon, I., Oliver, S., Salaman, D. M., Irwin, M., & Weatherley, S. J. 2004, *MNRAS*, **353**, 654
- Bahcall, J. N. & Peebles, P. J. E. 1969, *ApJ*, **156**, L7+
- Barkana, R. & Loeb, A. 2001, *Phys. Rep.*, **349**, 125
- Barlow, T. A., Junkkarinen, V. T., Burbidge, E. M., Weymann, R. J., Morris, S. L., & Korista, K. T. 1992, *ApJ*, **397**, 81
- Baum, W. A. 1962, in G. C. McVittie (ed.), *IAU Symp. 15 : Problems of Extra-Galactic Research*, pp 390–+
- Bazell, D. & Miller, D. J. 2005, *American Astronomical Society Meeting Abstracts*, **207**, 29.10
- Bechtold, J. 1994, *ApJS*, **91**, 1
- Becker, G. D., Sargent, W. L. W., & Rauch, M. 2004, *ApJ*, **613**, 61
- Bélanger, G., Goldwurm, A., Goldoni, P., & Paul, J. 2003, in F. Combes, D. Barret, T. Contini, & L. Paganì (eds.), *SF2A-2003 : Semaine de l'Astrophysique Française*, pp 377–+
- Bell, M. B. 2004, *ApJ*, **616**, 738
- Bergeron, P., Wesemael, F., & Beauchamp, A. 1995, *PASP*, **107**, 1047
- Bertelli, G., Bressan, A., Chiosi, C., Fagotto, F., & Nasi, E. 1994, *A&AS*, **106**, 275
- Bertin, E. & Arnouts, S. 1996, *A&AS*, **117**, 393
- Blumenthal, G. R., Faber, S. M., Primack, J. R., & Rees, M. J. 1984, *Nature*, **311**, 517
- Boksenberg, A., Sargent, W. L. W., & Rauch, M. 2003, *astro-ph/0307557*
- Boyle, B. J., Shanks, T., Croom, S. M., Smith, R. J., Miller, L., Loaring, N., & Heymans, C. 2000, *MNRAS*, **317**, 1014
- Bridle, A. H., Hough, D. H., Lonsdale, C. J., Burns, J. O., & Laing, R. A. 1994, *AJ*, **108**, 766
- Brotherton, M. S., Tran, H. D., Becker, R. H., Gregg, M. D., Laurent-Muehleisen, S. A., & White, R. L. 2001, *ApJ*, **546**, 775

- Carswell, R. F., Webb, J. K., Baldwin, J. A., & Atwood, B. 1987, *ApJ*, **319**, 709
- Cen, R., Miralda-Escude, J., Ostriker, J. P., & Rauch, M. 1994, *ApJ*, **437**, L9
- Charlton, J. C., Anninos, P., Zhang, Y., & Norman, M. L. 1997, *ApJ*, **485**, 26
- Chieze, J.-P., Alimi, J.-M., & Teyssier, R. 1998, *ApJ*, **495**, 630
- Choudhury, T. R., Padmanabhan, T., & Srianand, R. 2001, *MNRAS*, **322**, 561
- Claeskens, J.-F., Smette, A., Vandenbulcke, L., & Surdej, J. 2006, *MNRAS*, **367**, 879
- Comastri, A., Fiore, F., Vignali, C., Matt, G., Perola, G. C., & La Franca, F. 2001, *MNRAS*, **327**, 781
- Condon, J. J., Cotton, W. D., Greisen, E. W., Yin, Q. F., Perley, R. A., Taylor, G. B., & Broderick, J. J. 1998, *AJ*, **115**, 1693
- Coppolani, F., Petitjean, P., Stoehr, F., Rollinde, E., Pichon, C., Colombi, S., Haehnelt, M. G., Carswell, B., & Teyssier, R. 2006a, *To be published in The ESO messenger*
- Coppolani, F., Petitjean, P., Stoehr, F., Rollinde, E., Pichon, C., Colombi, S., Haehnelt, M. G., Carswell, B., & Teyssier, R. 2006b, *MNRAS*, **370**, 1804
- Coppolani, F., Petitjean, P., Stoehr, F., Rollinde, E., Pichon, C., Colombi, S., Haehnelt, M., Carswell, B., & Teyssier, R. 2005, in F. Casoli, T. Contini, J. M. Hameury, & L. Pagani (eds.), *SF2A-2005 : Semaine de l'Astrophysique Francaise*, pp 671–+
- Couchman, H. M. P. & Rees, M. J. 1986, *MNRAS*, **221**, 53
- Cristiani, S., D'Odorico, S., D'Odorico, V., Fontana, A., Giallongo, E., & Savaglio, S. 1997, *MNRAS*, **285**, 209
- Cristiani, S., D'Odorico, S., Fontana, A., Giallongo, E., & Savaglio, S. 1995, *MNRAS*, **273**, 1016
- Croft, R. A. C., Weinberg, D. H., Bolte, M., Burles, S., Hernquist, L., Katz, N., Kirkman, D., & Tytler, D. 2002, *ApJ*, **581**, 20
- Croft, R. A. C., Weinberg, D. H., Katz, N., & Hernquist, L. 1998, *ApJ*, **495**, 44
- Croom, S. M., Rhook, K., Corbett, E. A., Boyle, B. J., Netzer, H., Loaring, N. S., Miller, L., Outram, P. J., Shanks, T., & Smith, R. J. 2002, *MNRAS*, **337**, 275
- Croom, S. M., Smith, R. J., Boyle, B. J., Shanks, T., Miller, L., Outram, P. J., & Loaring, N. S. 2004, *MNRAS*, **349**, 1397
- Crotts, A. P. S. & Fang, Y. 1998, *ApJ*, **502**, 16
- Davé, R., Hernquist, L., Katz, N., & Weinberg, D. H. 1999, *ApJ*, **511**, 521
- Dinshaw, N. & Impey, C. D. 1996, *ApJ*, **458**, 73
- Dinshaw, N., Impey, C. D., Foltz, C. B., Weymann, R. J., & Chaffee, F. H. 1994, *ApJ*, **437**, L87
- Dinshaw, N., Weymann, R. J., Impey, C. D., Foltz, C. B., Morris, S. L., & Ake, T. 1997, *ApJ*, **491**, 45
- Dobrzycki, A. & Bechtold, J. 1991, *ApJ*, **377**, L69
- D'Odorico, V., Cristiani, S., D'Odorico, S., Fontana, A., Giallongo, E., & Shaver, P. 1998, *A&A*, **339**, 678
- D'Odorico, V., Petitjean, P., & Cristiani, S. 2002, *A&A*, **390**, 13
- Efstathiou, G., Bond, J. R., & White, S. D. M. 1992, *MNRAS*, **258**, 1P
- Elitzur, M., Nenkova, M., & Ivezić, Z. 2004, in S. Aalto, S. Huttemeister, & A. Pedlar (eds.), *ASP Conf. Ser. 320 : The Neutral ISM in Starburst Galaxies*, pp 242–+
- Elvis, M., Wilkes, B. J., & Tananbaum, H. 1985, *ApJ*, **292**, 357
- Fabian, A. C., Miniutti, G., Gallo, L., Boller, T., Tanaka, Y., Vaughan, S., & Ross, R. R. 2004, *MNRAS*, **353**, 1071
- Fan, X. 1999, *AJ*, **117**, 2528

- Fan, X., Strauss, M. A., Richards, G. T., Newman, J. A., Becker, R. H., Schneider, D. P., Gunn, J. E., Davis, M., White, R. L., Lupton, R. H., Anderson, Jr., J. E., Annis, J., Bahcall, N. A., Brunner, R. J., Csabai, I., Doi, M., Fukugita, M., Hennessy, G. S., Hindsley, R. B., Ivezić, Ž., Knapp, G. R., McKay, T. A., Munn, J. A., Pier, J. R., Szalay, A. S., & York, D. G. 2001, *AJ*, **121**, 31
- Fan, X., Strauss, M. A., Schneider, D. P., Gunn, J. E., Lupton, R. H., Anderson, S. F., Becker, R. H., Davis, M., Newman, J. A., Richards, G. T., White, R. L., & SDSS Collaboration 1999, *Bulletin of the American Astronomical Society*, **31**, 1517
- Fan, X.-H. & Chen, J.-S. 1993, *A&A*, **277**, L5+
- Fang, L. Z. 1991, *A&A*, **244**, 1
- Fang, Y., Duncan, R. C., Crots, A. P. S., & Bechtold, J. 1996, *ApJ*, **462**, 77
- Finlator, K., Ivezić, Ž., Fan, X., Strauss, M. A., Knapp, G. R., Lupton, R. H., Gunn, J. E., Rockosi, C. M., Anderson, J. E., Csabai, I., Hennessy, G. S., Hindsley, R. B., McKay, T. A., Nichol, R. C., Schneider, D. P., Smith, J. A., York, D. G., & the SDSS Collaboration 2000, *AJ*, **120**, 2615
- Fioc, M. & Rocca-Volmerange, B. 1997, *A&A*, **326**, 950
- Foltz, C. B., Chaffee, Jr., F. H., & Black, J. H. 1988, *ApJ*, **324**, 267
- Foltz, C. B., Hewett, P. C., Chaffee, F. H., & Hogan, C. J. 1993, *AJ*, **105**, 22
- Francis, P. J., Hewett, P. C., Foltz, C. B., Chaffee, F. H., Weymann, R. J., & Morris, S. L. 1991, *ApJ*, **373**, 465
- Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, *AJ*, **111**, 1748
- Fukugita, M., Shimasaku, K., & Ichikawa, T. 1995, *PASP*, **107**, 945
- Gavignaud, I., Bongiorno, A., Paltani, S., & et al. 2006, *ArXiv Astrophysics e-prints*
- Giacconi, R. & Gursky, H. 1965, *Space Science Reviews*, **4**, 151
- Glikman, E., Helfand, D. J., & White, R. L. 2006, *ApJ*, **640**, 579
- Gower, J. F. R., Scott, P. F., & Wills, D. 1967, *MmRAS*, **71**, 49
- Gunn, J. E. & Peterson, B. A. 1965, *ApJ*, **142**, 1633
- Haardt, F. & Madau, P. 1996, *ApJ*, **461**, 20
- Hall, P. B., Gunn, J. E., Knapp, G. R., Narayanan, V. K., Strauss, M. A., Anderson, S. F., vanden Berk, D. E., Heckman, T. M., Krolik, J. H., Tsvetanov, Z. I., Zheng, W., Richards, G. T., Schneider, D. P., Fan, X., York, D. G., Geballe, T. R., Davis, M., Becker, R. H., & Brunner, R. J. 2002, in D. M. Crenshaw, S. B. Kraemer, & I. M. George (eds.), *ASP Conf. Ser. 255 : Mass Outflow in Active Galactic Nuclei : New Perspectives*, pp 161–+
- Hall, P. B., Osmer, P. S., Green, R. F., Porter, A. C., & Warren, S. J. 1996, *ApJ*, **462**, 614
- Hamann, F., Korista, K. T., & Morris, S. L. 1993, *ApJ*, **415**, 541
- Hartwick, F. D. A. & Schade, D. 1990, *ARA&A*, **28**, 437
- Hasinger, G., Burg, R., Giacconi, R., Schmidt, M., Truemper, J., & Zamorani, G. 2003, *VizieR Online Data Catalog*, **332**, 90482
- Hatziminaoglou, E., Mathez, G., & Pelló, R. 2000, *A&A*, **359**, 9
- Hernquist, L., Katz, N., Weinberg, D. H., & Miralda-Escudé, J. 1996, *ApJ*, **457**, L51+
- Hewett, P. C. & Foltz, C. B. 2003, *AJ*, **125**, 1784
- Hook, I. M., McMahon, R. G., Boyle, B. J., & Irwin, M. J. 1994, *MNRAS*, **268**, 305
- Huchra, J. & Burg, R. 1992, *ApJ*, **393**, 90
- Hui, L. & Gnedin, N. Y. 1997, *MNRAS*, **292**, 27
- Hui, L., Stebbins, A., & Burles, S. 1999, *ApJ*, **511**, L5

- Impey, C. D., Foltz, C. B., Petry, C. E., Browne, I. W. A., & Patnaik, A. R. 1996, *ApJ*, **462**, L53+
- Jackson, J. D. 1975, *Classical electrodynamics*, 92/12/31, New York : Wiley, 1975, 2nd ed.
- Jakobsen, P., Perryman, M. A. C., di Serego Alighieri, S., Ulrich, M. H., & Macchetto, F. 1986, *ApJ*, **303**, L27
- Jannuzi, B. T., Bahcall, J. N., Bergeron, J., Boksenberg, A., Hartig, G. F., Kirhakos, S., Sargent, W. L. W., Savage, B. D., Schneider, D. P., Turnshek, D. A., Weymann, R. J., & Wolfe, A. M. 1998, *ApJS*, **118**, 1
- Jansen, F. A. & XMM Science Operations Team 2000, *Bulletin of the American Astronomical Society*, **32**, 724
- Kaiser, N. 1987, *MNRAS*, **227**, 1
- Kim, T.-S., Hu, E. M., Cowie, L. L., & Songaila, A. 1997, *AJ*, **114**, 1
- Koo, D. C., Kron, R. G., & Cudworth, K. M. 1986, *PASP*, **98**, 285
- Krisciunas, K. & Margon, B. 1998, *Bulletin of the American Astronomical Society*, **30**, 837
- Krolik, J. H. 1999, *Active galactic nuclei : from the central black hole to the galactic environment*, Active galactic nuclei : from the central black hole to the galactic environment / Julian H. Krolik. Princeton, N. J. : Princeton University Press, c1999.
- Kron, R. G. & Chiu, L.-T. G. 1981, *PASP*, **93**, 397
- Ledoux, C., Petitjean, P., & Srianand, R. 2006, *ApJ*, **640**, L25
- Lejeune, T., Cuisinier, F., & Buser, R. 1997, *A&AS*, **125**, 229
- Lenz, D. D., Newberg, J., Rosner, R., Richards, G. T., & Stoughton, C. 1998, *ApJS*, **119**, 121
- Lespine, Y. & Petitjean, P. 1997, *A&A*, **317**, 416
- Levshakov, S. A. & Varshalovich, D. A. 1985, *MNRAS*, **212**, 517
- Liebert, J., Dahn, C. C., & Monet, D. G. 1988, *ApJ*, **332**, 891
- Loh, E. D. & Spillar, E. J. 1986, *ApJ*, **303**, 154
- Low, F. J., Cutri, R. M., Huchra, J. P., & Kleinmann, S. G. 1988, *ApJ*, **327**, L41
- Lupton, R. H., Gunn, J. E., & Szalay, A. S. 1999, *AJ*, **118**, 1406
- Lynds, R. 1971, *ApJ*, **164**, L73+
- Madau, P. 1995, *ApJ*, **441**, 18
- Malkan, M. A. & Sargent, W. L. W. 1982, *ApJ*, **254**, 22
- McDonald, P. 2003, *ApJ*, **585**, 34
- McDonald, P. & Miralda-Escudé, J. 1999, *ApJ*, **518**, 24
- McDonald, P., Miralda-Escudé, J., Rauch, M., Sargent, W. L. W., Barlow, T. A., Cen, R., & Ostriker, J. P. 2000, *ApJ*, **543**, 1
- Melott, A. L. 1980, *ApJ*, **241**, 889
- Miralda-Escudé, J., Cen, R., Ostriker, J. P., & Rauch, M. 1996, *ApJ*, **471**, 582
- Moller, P. & Jakobsen, P. 1990, *A&A*, **228**, 299
- Muecket, J. P. & Mueller, V. 1987, *Ap&SS*, **139**, 163
- Newberg, H. J. & Yanny, B. 1997, *ApJS*, **113**, 89
- Nusser, A. 1999, *Monthly Notices of the Royal Astronomical Society*, **303(1)**, 179
- Oke, J. B. & Gunn, J. E. 1983, *ApJ*, **266**, 713
- Oke, J. B. & Korycansky, D. G. 1982, *ApJ*, **255**, 11
- Oppenheimer, B. D. & Davé, R. 2006, *ArXiv Astrophysics e-prints*
- Osmer, P. S., Kennefick, J. D., Hall, P. B., & Green, R. F. 1998, *ApJS*, **119**, 189
- Ostriker, J. P., Bajtlik, S., & Duncan, R. C. 1988, *ApJ*, **327**, L35
- Outram, P. J., Hoyle, F., & Shanks, T. 2001, *MNRAS*, **321**, 497

- Petitjean, P. 1997, in J. Bergeron (ed.), *The Early Universe with the VLT.*, pp 266–+
- Petitjean, P. 1999, *EDP Sciences*, 24
- Petitjean, P., Bergeron, J., Carswell, R. F., & Puget, J. L. 1993, *MNRAS*, **260**, 67
- Petitjean, P., Mueket, J. P., & Kates, R. E. 1995, *A&A*, **295**, L9
- Petitjean, P., Rauch, M., & Carswell, R. F. 1994, *A&A*, **291**, 29
- Petitjean, P., Surdej, J., Smette, A., Shaver, P., Muecket, J., & Remy, M. 1998, *A&A*, **334**, L45
- Pettini, M., Rix, S. A., Steidel, C. C., Adelberger, K. L., Hunt, M. P., & Shapley, A. E. 2002, *ApJ*, **569**, 742
- Pichon, C., Scannapieco, E., Aracil, B., Petitjean, P., Aubert, D., Bergeron, J., & Colombi, S. 2003, *ApJ*, **597**, L97
- Pichon, C., Vergely, J. L., Rollinde, E., Colombi, S., & Petitjean, P. 2001, *MNRAS*, **326**, 597
- Pickles, A. J. 1998, *VizieR Online Data Catalog*, **611**, 863
- Pilkington, J. D. H. & Scott, J. F. 1965, *MmRAS*, **69**, 183
- Prescott, M. K. M., Impey, C. D., Cool, R. J., & Scoville, N. Z. 2006, *ApJ*, **644**, 100
- Press, W. H., Rybicki, G. B., & Schneider, D. P. 1993, *ApJ*, **414**, 64
- Prochaska, J. X., Herbert-Fort, S., & Wolfe, A. M. 2005, *ApJ*, **635**, 123
- Rauch, M. 1998, *ARA&A*, **36**, 267
- Rauch, M., Becker, G. D., Viel, M., Sargent, W. L. W., Smette, A., Simcoe, R. A., Barlow, T. A., & Haehnelt, M. G. 2005, *ApJ*, **632**, 58
- Rauch, M. & Haehnelt, M. G. 1995, *MNRAS*, **275**, L76
- Rauch, M., Sargent, W. L. W., Womble, D. S., & Barlow, T. A. 1996, *ApJ*, **467**, L5+
- Rees, M. J. 1986, *MNRAS*, **218**, 25P
- Reichard, T. A., Richards, G. T., Hall, P. B., Schneider, D. P., Vanden Berk, D. E., Fan, X., York, D. G., Knapp, G. R., & Brinkmann, J. 2003, *AJ*, **126**, 2594
- Richards, G. T., Fan, X., Newberg, H. J., Strauss, M. A., Vanden Berk, D. E., Schneider, D. P., Yanny, B., Boucher, A., Burles, S., Frieman, J. A., Gunn, J. E., Hall, P. B., Ivezić, Ž., Kent, S., Loveday, J., Lupton, R. H., Rockosi, C. M., Schlegel, D. J., Stoughton, C., SubbaRao, M., & York, D. G. 2002, *AJ*, **123**, 2945
- Richards, G. T., Nichol, R. C., Gray, A. G., Brunner, R. J., Lupton, R. H., Vanden Berk, D. E., Chong, S. S., Weinstein, M. A., Schneider, D. P., Anderson, S. F., Munn, J. A., Harris, H. C., Strauss, M. A., Fan, X., Gunn, J. E., Ivezić, Ž., York, D. G., Brinkmann, J., & Moore, A. W. 2004, *ApJS*, **155**, 257
- Richards, G. T., Strauss, M. A., Fan, X., & et al. 2006, *AJ*, **131**, 2766
- Risaliti, G. & Elvis, M. 2004, *A Panchromatic View of AGN*, pp 187–+, ASSL Vol. 308 : Supermassive Black Holes in the Distant Universe
- Robin, A. C., Haywood, M., Ojha, D. K., Bienayme, O., & Creze, M. 1995, in P. C. van der Kruit & G. Gilmore (eds.), *IAU Symp. 164 : Stellar Populations*, pp 388–+
- Rollinde, E., Petitjean, P., & Pichon, C. 2001, *A&A*, **376**, 28
- Rollinde, E., Petitjean, P., Pichon, C., Colombi, S., Aracil, B., D’Odorico, V., & Haehnelt, M. G. 2003, *MNRAS*, **341**, 1279
- Rollinde, E., Srianand, R., Theuns, T., Petitjean, P., & Chand, H. 2005, *MNRAS*, **361**, 1015
- Rybicki, G. B. & Lightman, A. P. 1979, *Radiative processes in astrophysics*, New York, Wiley-Interscience, 1979. 393 p.
- Sandage, A. & Luyten, W. J. 1967, *ApJ*, **148**, 767



- Sandage, A. & Tammann, G. A. 1981, *A revised Shapley-Ames Catalog of bright galaxies*, Washington : Carnegie Institution, 1981, Preliminary version
- Sargent, W. L. W. 1980, *Phys. Scr.*, **21**, 753
- Sargent, W. L. W. 1988, in J. C. Blades, D. A. Turnshek, & C. A. Norman (eds.), *Proceedings of the QSO Absorption Line Meeting*, pp 1–10
- Sargent, W. L. W., Boksenberg, A., & Steidel, C. C. 1988, *ApJS*, **68**, 539
- Sargent, W. L. W., Young, P. J., Boksenberg, A., & Tytler, D. 1980, *ApJS*, **42**, 41
- Saslaw, W. C. & Zipoy, D. 1967, *Nature*, **216**, 976
- Scannapieco, E., Pichon, C., & Aracil, B. 2006, *astro-ph/0503001*
- Schaye, J., Aguirre, A., Kim, T.-S., Theuns, T., Rauch, M., & Sargent, W. L. W. 2003, *ApJ*, **596**, 768
- Schaye, J., Theuns, T., Rauch, M., Efstathiou, G., & Sargent, W. L. W. 2000, *MNRAS*, **318**, 817
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, **500**, 525
- Schneider, D. P., Fan, X., Hall, P. B., & et al. 2003, *AJ*, **126**, 2579
- Schneider, D. P., Hall, P. B., & et al. 2005, *VizieR Online Data Catalog*, **7243**, 0
- Shaver, P. A. & Robertson, J. G. 1983, *ApJ*, **268**, L57
- Shields, G. A. 1978, *BAAS*, **10**, 690
- Smette, A., Robertson, J. G., Shaver, P. A., Reimers, D., Wisotzki, L., & Koehler, T. 1995, *A&AS*, **113**, 199
- Smette, A., Surdej, J., Shaver, P. A., Foltz, C. B., Chaffee, F. H., Weymann, R. J., Williams, R. E., & Magain, P. 1992, *ApJ*, **389**, 39
- Songaila, A. 2001, *ApJ*, **561**, L153
- Songaila, A. & Cowie, L. L. 1996, *AJ*, **112**, 335
- Spergel, D. N., Verde, L., Peiris, H. V., Komatsu, E., Nolte, M. R., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Page, L., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L. 2003, *ApJS*, **148**, 175
- Srianand, R. & Petitjean, P. 2001, *A&A*, **373**, 816
- Srianand, R., Petitjean, P., & Ledoux, C. 2000, *Nature*, **408**, 931
- Stoughton, C., Lupton, R. H., Bernardi, M., & et al. 2002, *AJ*, **123**, 485
- Sutherland, R. S. & Dopita, M. A. 1993, *ApJS*, **88**, 253
- Szokoly, G. P., Bergeron, J., Hasinger, G., Lehmann, I., Kewley, L., Mainieri, V., Nordin, M., Rosati, P., Giacconi, R., Gilli, R., Gilmozzi, R., Norman, C., Romaniello, M., Schreier, E., Tozzi, P., Wang, J. X., Zheng, W., & Zirm, A. 2004, *ApJS*, **155**, 271
- Tananbaum, H., Avni, Y., Branduardi, G., Elvis, M., Fabbiano, G., Feigelson, E., Giacconi, R., Henry, J. P., Pye, J. P., Soltan, A., & Zamorani, G. 1979, *ApJ*, **234**, L9
- Tegmark, M., Silk, J., Rees, M. J., Blanchard, A., Abel, T., & Palla, F. 1997, *ApJ*, **474**, 1
- Telfer, R. C., Zheng, W., Kriss, G. A., & Davidsen, A. F. 2002, *ApJ*, **565**, 773
- Teplitz, H. I., Malkan, M., & McLean, I. S. 1998, *ApJ*, **506**, 519
- Teyssier, R., Chièze, J.-P., & Alimi, J.-M. 1998, *ApJ*, **509**, 62
- Theuns, T., Leonard, A., Efstathiou, G., Pearce, F. R., & Thomas, P. A. 1998, *MNRAS*, **301**, 478
- Tolea, A., Krolik, J. H., & Tsvetanov, Z. 2002, *ApJ*, **578**, L31
- Trevese, D., Giallongo, E., & Camurani, L. 1992, *ApJ*, **398**, 491
- Tytler, D. 1982, *Nature*, **298**, 427
- Tytler, D. & Fan, X.-M. 1994, *ApJ*, **424**, L87
- Vanden Berk, D. E., Richards, G. T., & et al. 2001, *AJ*, **122**, 549

- Vanden Berk, D. E., Schneider, D. P., Richards, G. T., Hall, P. B., Strauss, M. A., Brunner, R., Fan, X., Baldry, I. K., York, D. G., Gunn, J. E., Nichol, R. C., Meiksin, A., & Brinkmann, J. 2005, *AJ*, **129**, 2047
- Viel, M., Matarrese, S., Mo, H. J., Haehnelt, M. G., & Theuns, T. 2002, *MNRAS*, **329**, 848
- Viel, M., Weller, J., & Haehnelt, M. G. 2004, *MNRAS*, **355**, L23
- VonNeumann, J. & Richtmyer, R. D. 1950, *Journal of Applied Physics*, **21**(3), 232
- Warren, S. J., Hewett, P. C., & Osmer, P. S. 1994, *ApJ*, **421**, 412
- Weinstein, M. A., Richards, G. T., Schneider, D. P., Younger, J. D., Strauss, M. A., Hall, P. B., Budavári, T., Gunn, J. E., York, D. G., & Brinkmann, J. 2004, *ApJS*, **155**, 243
- Weisskopf, M. C., Tananbaum, H. D., Van Speybroeck, L. P., & O'Dell, S. L. 2000, in J. E. Truemper & B. Aschenbach (eds.), *Proc. SPIE Vol. 4012, p. 2-16, X-Ray Optics, Instruments, and Missions III, Joachim E. Truemper; Bernd Aschenbach; Eds.*, pp 2–16
- Weymann, R. J., Carswell, R. F., & Smith, M. G. 1981, *ARA&A*, **19**, 41
- Weymann, R. J., Morris, S. L., Foltz, C. B., & Hewett, P. C. 1991, *ApJ*, **373**, 23
- Williger, G. M., Smette, A., Hazard, C., Baldwin, J. A., & McMahon, R. G. 2000, *ApJ*, **532**, 77
- Willott, C. J., Delfosse, X., Forveille, T., Delorme, P., & Gwyn, S. D. J. 2005, *ApJ*, **633**, 630
- Wolf, C., Meisenheimer, K., Kleinheinrich, M., Borch, A., Dye, S., Gray, M., Wisotzki, L., Bell, E. F., Rix, H.-W., Cimatti, A., Hasinger, G., & Szokoly, G. 2004, *A&A*, **421**, 913
- Wolf, C., Meisenheimer, K., & Röser, H.-J. 2001, *A&A*, **365**, 660
- Wolfe, A. M., Turnshek, D. A., Smith, H. E., & Cohen, R. D. 1986, *ApJS*, **61**, 249
- Worsley, M. A., Fabian, A. C., Barcons, X., Mateos, S., Hasinger, G., & Brunner, H. 2004, *MNRAS*, **352**, L28
- Xia, L., Zhou, X., Ma, J., Wu, H., Sun, W.-H., Jiang, Z., Xue, S., Chen, J., & Chen, W. 2002, *PASP*, **114**, 1349
- Yip, C. W., Connolly, A. J., Vanden Berk, D. E., Ma, Z., Frieman, J. A., SubbaRao, M., Szalay, A. S., Richards, G. T., Hall, P. B., Schneider, D. P., Hopkins, A. M., Trump, J., & Brinkmann, J. 2004, *AJ*, **128**, 2603
- Young, P., Sargent, W. L. W., & Boksenberg, A. 1982, *ApJ*, **252**, 10
- Young, P. A., Impey, C. D., & Foltz, C. B. 2001, *ApJ*, **549**, 76
- Zhang, Y., Anninos, P., & Norman, M. L. 1995, *ApJ*, **453**, L57+
- Zuo, L. & Bond, J. R. 1994, *ApJ*, **423**, 73